

Hart, Edward

89452

From: Ungar, Susan
Sent: Wednesday, March 19, 2003 3:30 PM
To: Hutzell, Paula
Cc: Hart, Edward
Subject: Rush Sequence Search for 09/674,975

Hi

I need a rush sequence search for 09/674,975, SEQ ID NOs 1 and 2. These are both peptides.
Just realized that I need to get this case in this biweek.

Please forward this search to Ed Hart who has agreed to submit it for me.

Thanks
Susan Ungar
1642
703-305-2181
CMI-8B05

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

SEI2
MB

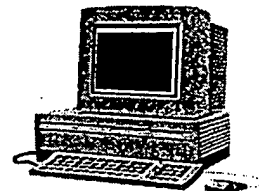
3/19/03
Compurgan/P
2-AA

3/20/03
completed

BioTech-Chem Library

Search Results

Feedback Form (Optional)



Scientific & Technical Information

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech-searcher* who conducted the search *or* contact:

Mary Hale, Supervisor, 308-
CM-1 Room 1E01

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:* (Example: 1610)

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

Other Comments:

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or mary.hale@uspto.gov

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2003, 17:18:13 ; Search time 32.5 Seconds

(without alignments)
180.401 Million cell updates/sec

Title: US-09-674-975-1

Perfect score: 243
Sequence: 1 KISHFLKMSLNFIKRAHPY.....NCEPANPSEKNSPSTGYCS 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	243	100.0	44	21	AA55000	Human CD20 protein
2	243	100.0	297	10	AA91356	CD20.4 antigen.
3	243	100.0	297	13	AA20808	Human CD20 antigen
4	243	100.0	297	17	AA91436	Human CD20 antigen
5	243	100.0	297	19	AA80445	Human CD20.4 antigen
6	243	100.0	297	20	AA86192	Human CD20.4 antigen
7	243	100.0	297	21	AA96131	Human cell surface
8	243	100.0	297	22	AA02440	Human lymphocyte c
9	212	87.2	150	22	AA80564	Human haematologic
10	212	87.2	150	22	AA80569	Human haematologic

11	212	87.2	150	22	AA80610	Human haematologic
12	212	87.2	150	22	AA80619	Human haematologic
13	212	87.2	150	22	AA80950	Human haematologic
14	212	87.2	150	22	AA81576	Human haematologic
15	196	80.7	146	22	AA80468	Human haematologic
16	184	75.7	94	21	AA80896	Human haematologic
17	160	65.8	44	21	AA55001	Mouse CD20 protein
18	144	59.3	137	22	AA80956	Human haematologic
19	97	39.9	141	22	AA80456	Human haematologic
20	61.5	25.3	495	20	AA55997	Human endometrium
21	61.5	25.3	750	20	AA835328	Human TR13 receptor
22	61.5	25.3	911	22	AA83851	Amino acid sequenc
23	61.5	25.3	911	22	AA83850	Human TR13 receptor
24	61.5	25.3	1001	22	AA835333	Human CAS8619 prot
25	61.5	25.3	1013	21	AA826179	Human PRO4985 poly
26	61.5	25.3	1013	22	AA012190	Human acid sequenc
27	61.5	25.3	1013	22	AA83845	Cardamoyl-phosphat
28	60.5	24.9	2391	15	AA85664	Novel human secret
29	56.5	23.3	160	22	AA029533	Human histocompati
30	56	23.0	10	21	AA55002	Human polypeptide
31	56	23.0	48	22	AA010583	Staphylococcus epi
32	56	23.0	101	23	ABP40241	Hericidially activ
33	54.5	22.4	347	23	AB81886	Sequence of murine
34	54.5	22.4	556	15	AA831355	TR16-short receptor
35	54.5	22.4	963	22	AA870255	Novel human diagno
36	54.5	22.4	1027	22	AA870256	Novel human diagno
37	54	22.2	445	22	AA08252	G protein-coupled
38	54	22.2	447	20	AA135554	Novel human secret
39	53.5	22.0	393	17	AA89597	C. pneumoniae prot
40	53	21.8	142	20	AA48286	Chimeric protein M
41	53	21.8	348	22	ABG26386	Human prostate can
42	53	21.8	505	22	ABG27093	Novel human diagno
43	53	21.8	1132	22	ABG07312	Novel human diagno
44	52.5	21.6	177	22	AA025634	Novel human secret
45	52	21.6	115	22	AA031606	Novel human secret

ALIGNMENTS

```

RESULT 1
AA55000
AA55000 standard; peptide: 44 AA.
ID
XX
AC AA55000;
AC
XX 18-FEB-2000 (first entry)
DT
XX
XX Human CD20 protein fragment.
DE
XX
XX CD20; active vaccination; autologous cell; transmembrane protein; cancer;
KW B cell Non-Hodgkin's lymphoma; B cell disease; autoimmune disorder;
KW chronic lymphocytic leukaemia; B-cell regulatory disorder; therapy;
human.
XX
XX Homo sapiens.
OS
XX
XX
XX
XX W09957981-A1.
XX
XX 18-NOV-1999.
PD
XX
XX 07-MAY-1999; 99WO-US10065.
PF
XX
XX 08-MAY-1998; 98US-0084870.
PR
XX
XX (SLOC ) SLOAN KETTERING INST CANCER RES.
PA
XX
XX Agus DB, Scheinberg D, Roberts W, Zelenetz AD;
PI
XX WPI: 2000-038920/03.
XX
XX
XX Novel compositions and methods utilizing a humoral immune response to
XX treat B cell Non-Hodgkin's lymphoma and other diseases
PT

```

XX Claim 4; Page 27; 43pp; English.
 PS
 CC This sequence represents a fragment of the human CD20 protein. The
 CC invention relates to a method for active vaccination against autologous
 CC cells expressing transmembrane proteins comprising administering to a
 CC patient a vaccine composition comprising at least an immunogenic portion
 CC of the extracellular domain of the transmembrane protein, or a xenogeneic
 CC homolog thereof, coupled to or administered with a carrier protein
 CC effective to break tolerance to the transmembrane protein and a
 CC pharmaceutically acceptable adjuvant. The method of the invention is used
 CC to treat a variety of cancers, especially B cell Non-Hodgkin's lymphoma
 CC (NHL) and other B cell diseases such as chronic lymphocytic leukaemia,
 CC autoimmune disorders and B-cell regulatory disorders. The method is also
 CC useful for treating diseases and conditions where administration of
 CC antibodies to transmembrane proteins has been shown to provide
 CC therapeutic benefits, and especially where the target is also capable of
 CC transducing or receiving a signal important for cell growth or function.
 CC While therapeutic regimens using anti-CD20 concepts are potentially
 CC effective for the treatment of B cell Non-Hodgkin's lymphoma, they are
 CC passive therapies that do not directly involve the immune system of the
 CC patient. Thus, these therapies may require continued administration of
 CC the therapeutic agent for efficacy, and do not provide any long term
 CC protection against recurrence. The present invention uses a humoral
 CC immune response against CD20 found in B cells. The method does not need
 CC long term administration of the therapeutic agent, and provides long term
 CC protection.
 CC
 CC Sequence 44 AA:
 SQ
 Query Match 100.0%; Score 243; DB 21; Length 44;
 Best Local Similarity 100.0%; Pred. No. 3e-25;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 KISHFLKMSLNFRHPTPIYINYNCEPANPSEKNSPTQYCYCS 44
 1 KISHFLKMSLNFRHPTPIYINYNCEPANPSEKNSPTQYCYCS 44
 Db
 RESULT 2
 AAP91356
 ID AAP91356 standard; protein; 297 AA.
 XX
 AC AAP91356;
 XX
 DT 20-DEC-1989 (first entry).
 XX
 DE CD20.4 antigen.
 XX
 KW Tyrosine suppressor gene; ASV LTR; human cytomegalovirus AD169 enhancer;
 KW HIV box; immunoselection; immune deficiency diseases; vasculitis;
 KW systemic lupus erythematosus; rheumatoid arthritis; neoplasms;
 XX
 FH Key
 FT 9..11 Location/Qualifiers
 FT Binding-site
 FT /note= N-linked glycosylation
 FT 293..295
 FT /note= N-linked glycosylation
 FT 51..103
 FT /note= hydrophobic
 FT 117..141
 FT /note= hydrophobic
 FT 183..203
 FT /note= hydrophobic
 FT Region
 FT 183..203
 FT /note= hydrophobic
 XX
 XX EP330191-A.
 XX
 PD 30-AUG-1989.
 XX
 XX 23-FEB-1989; 89EP-0103127.
 XX
 XX 25-FEB-1988; 88US-0160416.
 XX

PA (GCHO) GEN. HOSPITAL CORP.
 XX
 PI Seed B, Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo CP;
 PI Simmons D, Stamenkovic I, Stengelin S;
 XX
 DR WPI: 1989-250302/35.
 DR N-PSDB: AAN90610.
 XX
 PS Rapid immuno:selection cloning - used to clone genes encoding
 PT cell surface antigens associated with mammalian T lymphocytes.
 XX
 CC Disclosure; Fig 10; 69pp; English.
 XX
 CC CD20.4 cell surface antigen involved in cell mediated immunity.
 CC Its DNA can be expressed in a vector which transforms COS cells. The
 CC vector can isolate any protein, and clones are easy to manipulate.
 XX
 SQ Sequence 297 AA:
 Query Match 100.0%; Score 243; DB 10; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 KISHFLKMSLNFRHPTPIYINYNCEPANPSEKNSPTQYCYCS 44
 142 KISHFLKMSLNFRHPTPIYINYNCEPANPSEKNSPTQYCYCS 185
 Db
 RESULT 3
 AAR20808
 ID AAR20808 standard; protein; 297 AA.
 XX
 AC AAR20808;
 XX
 DT 21-MAY-1992 (first entry)
 XX
 DE Human CD20 antigen.
 XX
 KW Cloning technique; cell surface antigen; immunodiagnosis;
 KW B cell activation; Fc receptor; Immunoglobulin G; IgG; tumour.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT 51..103 Location/Qualifiers
 FT Region
 FT /label= hydrophobic
 FT /note= "possibly acts as internal membrane
 FT insertion signal"
 FT 117..141
 FT /label= hydrophobic
 FT 183..203
 FT /label= hydrophobic
 FT 9..11
 FT /label= N-linked glycosylation.site
 FT /note= "probably not used if located
 FT intracellularly"
 FT Modified-site
 FT 293..295
 FT /label= N-linked-glycosylation.site
 FT /note= "probably not used if located
 FT intracellularly"
 XX
 XX WO9201049-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90WO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GCHO-) GEN HOSPITAL CORP.
 XX
 XX Seed B, Aruffo A, Amlot M;
 XX

DR WPI; 1992-056864/07.
 DR N-PSDB; AAQ21170.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 PS Example 6; Fig 10; 160pp; English.
 CC Clone CD20.4 was one of two cDNA clones to be isolated from a human
 CC Burkitt cell line Daudi library by panning with a panel of MAbs
 CC against CD20. The amino acid sequence predicted from the nucleotide
 CC sequence has a highly charged carboxy-terminal domain and lacks an
 CC amino-terminal signal peptide. The two putative N-linked
 CC glycosylation sites are not expected to be used so the difference
 CC in Mr between CD20 expressed on COS cells and on B cells is
 CC presumed to be due to O-linked glycosylation. The proposed
 CC extracellular domain (residues 142 to 182) is rich in Ser and Thr
 CC residues which might support O-glycosylation. Given that the
 CC greater part of the protein is apparently confined to the membrane
 CC or the cytoplasm, it is suggested that CD20 may be involved in
 CC signal transduction from other transmembrane proteins to the cell
 CC interior.
 CC
 XX
 SQ Sequence 297 AA:
 Query Match 100.0%; Score 243; DB 13; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KISHFLKMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 44
 ||||||||||||||||||||||||||||||||||||||||
 Db 142 KISHFLKMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 185
 RESULT 4
 AAR91436
 ID AAR91436 standard; Protein; 297 AA.
 XX
 AC AAR91436;
 DT 30-OCT-1996 (first entry)
 XX
 DE Human CD20 antigen.
 XX
 KM Cell surface antigen; cloning; immunoselection; immunotherapy;
 KM therapy; diagnosis; vector; CD20; COS; B-lymphocyte.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 9..12
 FT /label= N-glycosylation_site
 FT Domain 512..103
 FT /label= Hydrophobic_domain
 FT Domain 117..141
 FT /label= Hydrophobic_domain
 FT Domain 183..203
 FT /label= Hydrophobic_domain
 FT Modified-site 293..295
 FT /label= N-glycosylation_site
 FT
 XX
 XX US5506126-A.
 PN
 XX
 PD 09-APR-1996.
 XX
 PF 25-FEB-1988; 880S-0160416.
 XX
 PR 01-DEC-1992; 920S-0983647.
 PR 25-FEB-1988; 880S-0160416.
 PR 13-JUL-1989; 890S-0379076.
 PR 13-JUL-1990; 900S-0553759.
 PR 18-OCT-1993; 930S-0139273.
 XX

PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Arufio A, Seed B;
 XX
 DR WPI; 1996-200279/20.
 DR N-PSDB; AAT14710.
 XX
 PT Cloning of cDNA encoding cell surface antigen - useful for isolation
 PT of diagnostic and therapeutic proteins
 XX
 PS Example 6; Fig 10A-B; 79pp; English.
 XX
 XX The amino acid sequence (AAR91436) of CD20, a human pan B-cell
 CC antigen, was deduced from cDNA clone CD20.4 (AAT14710) derived from
 CC Daudi cells. CD20 was expressed in COS cells following construction
 CC of a cDNA library and panning of the library using a panel of
 CC monoclonal antibodies against CD20. Using the novel immunoselection
 CC cloning method, cell surface antigens (see also AAR91431-46) can be
 CC obtd. for diagnostic and therapeutic use in cases of immune-
 CC associated disease, and for identification, isolation and
 CC purification of antibodies and antigens.
 CC
 XX
 SQ Sequence 297 AA:
 Query Match 100.0%; Score 243; DB 17; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KISHFLKMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 44
 ||||||||||||||||||||||||||||||||||||||||
 Db 142 KISHFLKMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 185
 RESULT 5
 AAW80445
 ID AAW80445 standard; Protein; 297 AA.
 XX
 AC AAW80445;
 DT 07-JUN-1999 (first entry)
 XX
 DE Human CD20.4 antigen.
 XX
 KM CD20.4; cell surface antigen; human; B lymphocyte.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Modified-site 9..11
 FT /note= "Asn is N-glycosylated"
 FT Region 51..103
 FT /note= "hydrophobic region"
 FT Region 117..141
 FT /note= "hydrophobic region"
 FT Region 183..203
 FT /note= "hydrophobic region"
 FT Modified-site 293..295
 FT /note= "Asn is N-glycosylated"
 FT
 XX
 XX US5830731-A.
 PN
 XX
 PD 03-NOV-1998.
 XX
 PF 21-MAY-1997; 970S-0861205.
 XX
 PR 01-DEC-1992; 920S-0983647.
 PR 25-FEB-1988; 880S-0160416.
 PR 13-JUL-1989; 890S-0379076.
 PR 13-JUL-1990; 900S-0553759.
 PR 21-MAY-1997; 970S-0861205.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 XX

PI Aruffo A, Seed B;
 XX
 DR WPI: 1998-609251/51.
 DR N-PSDB: AAV63448.
 XX
 PT New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX
 PS Example 6; Fig 10A-B; 79pp; English.

CC This polypeptide comprises human CD20.4 antigen. Its amino acid
 CC sequence was deduced from the nucleotide sequence (see AAV63448) of a
 CC cDNA clone isolated from a Daudi cell cDNA library using a novel
 CC method for cloning cDNAs from mammalian expression libraries. The
 CC method is based on transient expression of an antigen in eukaryotic
 CC cells and physical selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. It is useful for the
 CC isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell, and
 CC has been used to clone genes (see AAV63442-63) encoding cell surface
 CC antigens from mammalian lymphocytes (see AAV80440-55). CD20, a
 CC B-lymphocyte specific antigen (B1, Bp35), was expressed in COS
 CC cells as 2 protein species, CD20.4 and CD20.6. The purified genes
 CC and proteins are useful for immunodiagnosis and immunotherapeutic
 CC applications, including the diagnosis and treatment of
 CC immune-mediated infections, diseases, and disorders of animals,
 CC including humans.
 CC
 CC Sequence 297 AA;

Query Match 100.0%; Score 243; DB 19; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISHFLKMSLNFIKRAHPYININCEPANPSEKNSPSTQCYCS 44
 ||||||||||||||||||||||||||||||||||||||||
 DB 142 KISHFLKMSLNFIKRAHPYININCEPANPSEKNSPSTQCYCS 185

RESULT 6
 AAV86192
 ID AAV86192 standard; Protein: 297 AA.
 XX
 AC AAV86192;

DE 10-MAY-1999 (first entry)

DE Human CD20.4 antigen.

KM CD20.4; cell surface antigen; human; cDNA library; B lymphocyte.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Modified-site /note="Asn is N-glycosylated"

FT Region 51..103

FT Region /note="hydrophobic region"

FT Region 117..141

FT Region /note="hydrophobic region"

FT Region 183..203

FT Modified-site /note="hydrophobic region"

FT 293..295

PR 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0498809.
 PR 13-JUL-1990; 90US-0553759.
 PR 07-JUN-1995; 95US-0485447.

XX (GENO) GEN HOSPITAL CORP.

PI Allen J, Amlot M, Aruffo A, Camerini D, Laufer L,
 PI Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;
 DR WPI: 1999-069813/06.
 DR N-PSDB: AAV81206.

PT cDNA encoding human CD40 antigen - useful for cloning cDNA encoding
 PT cell surface antigens, constructing cDNA libraries, expression
 PT vectors for expression in eukaryotic cells or their fragments
 XX
 PS Example 6; Fig 10A-B; 79pp; English.

CC This polypeptide comprises human CD20.4 antigen. Its amino acid
 CC sequence was deduced from the nucleotide sequence (see AAV81205) of a
 CC cDNA clone isolated from a Daudi cell cDNA library using a novel
 CC method for cloning cDNAs from mammalian expression libraries. The
 CC method is based on transient expression of an antigen in eukaryotic
 CC cells and physical selection of cells expressing the antigen by
 CC adhesion to an antibody-coated substrate. It is useful for the
 CC isolation and cloning of any protein which can be expressed and
 CC transported to the cell surface membrane of a eukaryotic cell, and
 CC has been used to clone genes (see AAV81198-220) encoding cell surface
 CC antigens such as CD1a, CD1b, CD1c, CD2, CD6, CD7, CD13, CD14, CD16,
 CC CD19, CD20, CD22, CD26, CD27, CD28, CD31, CD32a, CD32b, CD33,
 CC CD34, CD36, CD37, CD38, CD39, CD40, CD43, CD44, CD53, ICAM, LFA-3,
 CC FCRII, FCRIIb, T11sa and LFA-1 (see AAV86188-62, AAV89151-52 and
 CC AAV88451). CD40 cDNA (see AAV81198) is specifically claimed. CD20, a
 CC B-lymphocyte specific antigen (B1, Bp35), was expressed in COS
 CC cells as 2 protein species, CD20.4 and CD20.6.
 CC
 CC Sequence 297 AA;

Query Match 100.0%; Score 243; DB 20; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISHFLKMSLNFIKRAHPYININCEPANPSEKNSPSTQCYCS 44
 ||||||||||||||||||||||||||||||||||||||||
 DB 142 KISHFLKMSLNFIKRAHPYININCEPANPSEKNSPSTQCYCS 185

RESULT 7

AAV96131
 ID AAV96131 standard; Protein: 297 AA.

AC AAV96131;

DE 19-DEC-2000 (first entry)

DE Human cell surface antigen CD20.4.

KM CD20.4; cell surface antigen; human; immunoselection; panning;
 KM immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 KW immune disorder; infection; asthma; immune-complex disease;
 KW amyloidosis; multiple sclerosis; parasitic disease.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Region 51..103

FT Region /note="hydrophobic region"

FT Region 117..141

FT Region /note="hydrophobic region"

FT 183..203

FT Modified-site /note="hydrophobic region"

FT Modified-site 9 /note="N-glycosylated"
 FT Modified-site 293 /note="N-glycosylated"
 FT
 XX US611093-A.
 XX
 PD 29-AUG-2000.
 XX
 PF 28-OCT-1998; 98US-0181612.
 XX
 PR 01-DEC-1992; 92US-0983647.
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 23-MAR-1990; 90US-0498809.
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Stamenkovic I, Seed B;
 XX
 DR WPI: 2000-586382/55.
 DR N-PSDB; AAS0584.
 XX
 XX
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases
 XX
 PS Example 6; Fig 10A-B; 75pp; English.
 XX
 CC The present sequence is that of human cell surface antigen (CSA)
 CC CD20.4, as predicted from cDNA isolated from a human Burkitt cell
 CC line Daudi cell library. CD20.4 (or B1, Bp35) is a pan B-cell
 CC antigen that plays an important role in B cell activation. The
 CC cDNA (see AAS0584) was identified using a new method for cloning
 CC cDNAs encoding CSAs. The method is based upon transient expression
 CC of CSA in eukaryotic cells and physical selection of cells expressing
 CC the CSA by adhesion to (panning on) an antibody-coated substrate
 CC such as a culture dish. CSA nucleic acids isolated by the method
 CC of the invention, and the proteins they encode, are useful for
 CC immunodiagnostic and immunotherapeutic applications, including the
 CC diagnosis and treatment of immune-mediated infections, diseases,
 CC and disorders in animals, including humans. These disorders
 CC include asthma, immune-complex disease, amyloidosis, parasitic
 CC diseases or multiple sclerosis.
 CC
 SQ Sequence 297 AA:
 Query Match 100.0%; Score 243; DB 21; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2,9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KISFLKMSINLIRAHPTPIYINCEPANSEKNSPTQYCS 44
 Db 142 KISFLKMSINLIRAHPTPIYINCEPANSEKNSPTQYCS 185
 RESULT 8
 AAU02440
 ID AAU02440 standard; Protein; 297 AA.
 XX
 AC AAU02440;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Human lymphocyte cell surface antigen CD20 polypeptide.
 XX
 KM Human: lymphocyte cell surface antigen; immune-mediated disease; CD20;
 KM infection; immune deficiency disorder; hypersensitivity; inflammation;
 KM systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;
 KM transplant rejection; asthma.
 XX

OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 9..11
 FT Region /note="Asn is glycosylated"
 FT Region 51..103
 FT Region /note="Hydrophobic region"
 FT Region 117..141
 FT Region /note="Hydrophobic region"
 FT Region 183..203
 FT Region /note="Hydrophobic region"
 FT Modified-site 293..295
 FT /note="Asn is glycosylated"
 XX
 XX US6218525-B1.
 XX
 PD 17-APR-2001.
 XX
 PF 01-DEC-1992; 92US-0983647.
 XX
 PR 25-FEB-1988; 88US-0160416.
 PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Simmons D;
 XX
 DR WPI: 2001-289848/30.
 DR N-PSDB; AAS03178.
 XX
 XX
 PT New recombinant DNA encoding CD28 useful for diagnosing and treating
 PT immune-mediated diseases, infections or disorders, e.g. systemic lupus
 PT erythematosus, asthma, transplant rejection, rheumatoid arthritis
 PT
 XX
 PS Example 6; Fig 10A-10B; 72pp; English.
 XX
 CC The present sequence representing human lymphocyte cell surface antigen
 CC CD20 is 1 of various human lymphocyte cell surface antigen
 CC polypeptide sequences (AAU02435-AAU02432) described in the present
 CC invention. The invention relates to a novel method of cloning cDNA
 CC encoding cell surface antigens and efficient construction of cDNA
 CC libraries. Also described are 2 expression vectors (AAS03171, AAS03174)
 CC which provide high level expression in eukaryotic host cells. A
 CC genetically engineered cDNA sequence encoding the CD28 amino acid
 CC extracellular domain sequence (amino acids 1-134 given in AAU02437)
 CC and/or comprising nucleotides 100-759, 134-555 or 134-759 of the CD28
 CC cDNA sequence (AAS03175) is also new. The purified genes and proteins
 CC are useful for immunodiagnostic and immunotherapeutic applications, such
 CC as in the diagnosis and treatment of immune-mediated diseases, infections
 CC or disorders in animals and humans. Such diseases include immune
 CC deficiency diseases, diseases of immediate type of hypersensitivity,
 CC asthma, hypersensitivity pneumonitis, systemic lupus erythematosus,
 CC rheumatoid arthritis, acute and chronic inflammation, platelet disorders,
 CC plasma and other cell neoplasms, parasitic diseases, multiple sclerosis,
 CC Guillain-Barre syndrome and tissue and organ transplant rejection. The
 CC sequences can also be used to identify, isolate and purify other
 CC antibodies and antigens.
 CC
 SQ Sequence 297 AA:
 Query Match 100.0%; Score 243; DB 22; Length 297;
 Best Local Similarity 100.0%; Pred. No. 2,9e-24;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KISFLKMSINLIRAHPTPIYINCEPANSEKNSPTQYCS 44
 Db 142 KISFLKMSINLIRAHPTPIYINCEPANSEKNSPTQYCS 185
 RESULT 9
 AAM80564
 ID AAM80564 standard; Protein; 150 AA.

```

XX AC AAM80564;
XX
XX DT 13-NOV-2001 (first entry)
XX DE Human haematological malignancy-related antigen #262.
XX DE
XX DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
XX OS Homo sapiens.
XX PN WO200164886-A2.
XX PD
XX PD 07-SEP-2001.
XX PF 01-MAR-2001; 2001WO-US07272.
XX PR 01-MAR-2000; 2000US-0186126.
XX PR 17-MAR-2000; 2000US-0190479.
XX PR 27-APR-2000; 2000US-0200545.
XX PR 28-APR-2000; 2000US-0200303.
XX PR 28-APR-2000; 2000US-0200779.
XX PR 01-MAY-2000; 2000US-0200999.
XX PR 04-MAY-2000; 2000US-0202084.
XX PR 12-MAY-2000; 2000US-0206201.
XX PR 14-JUL-2000; 2000US-0218950.
XX PR 03-AUG-2000; 2000US-0222903.
XX PR 04-AUG-2000; 2000US-0223416.
XX PR 07-AUG-2000; 2000US-0223378.
XX PA (CORI-) CORIXA CORP.
XX PI
XX PI Galger A, Algate PA, Mannion J;
XX DR WPI: 2001-514842/56.
XX PT Compositions and methods for the detection of hematological
XX PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX PS Claim 1; Page 617; 1252pp; English.
XX CC The present invention relates to compositions and methods for the
XX CC detection, diagnosis and therapy of haematological malignancies. The
XX CC present sequence is the protein sequence of a human haematological
XX CC malignancy related antigen. The methods of the present invention comprise
XX CC detecting the presence of haematological malignancy related antigen(s) in
XX CC a sample obtained from the patient (an increased level of the
XX CC polypeptide, compared to an unaffected individual, is indicative of an
XX CC increased risk). Haematological malignancies which can be treated using
XX CC the present invention are chronic lymphocytic leukaemia, lymphoma,
XX CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX CC cell non-Hodgkin's lymphoma.
XX SO Sequence 150 AA:
XX
XX Query Match 87.2%; Score 212; DB 22; Length 150;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-20;
XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 KMESLNFI RAHTPYININCEPANSEKNSPSTQCYCS 44
DB 1 KMESLNFI RAHTPYININCEPANSEKNSPSTQCYCS 38

```

```

DT DE 13-NOV-2001 (first entry)
XX DE
XX DE Human haematological malignancy-related antigen #267.
XX DE
XX DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;
XX KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
XX OS Homo sapiens.
XX PN WO200164886-A2.
XX PD
XX PD 07-SEP-2001.
XX PF 01-MAR-2001; 2001WO-US07272.
XX PR 01-MAR-2000; 2000US-0186126.
XX PR 17-MAR-2000; 2000US-0190479.
XX PR 27-APR-2000; 2000US-0200545.
XX PR 28-APR-2000; 2000US-0200303.
XX PR 28-APR-2000; 2000US-0200779.
XX PR 01-MAY-2000; 2000US-0200999.
XX PR 04-MAY-2000; 2000US-0202084.
XX PR 12-MAY-2000; 2000US-0206201.
XX PR 14-JUL-2000; 2000US-0218950.
XX PR 03-AUG-2000; 2000US-0222903.
XX PR 04-AUG-2000; 2000US-0223416.
XX PR 07-AUG-2000; 2000US-0223378.
XX PA (CORI-) CORIXA CORP.
XX PI
XX PI Galger A, Algate PA, Mannion J;
XX DR WPI: 2001-514842/56;
XX PT Compositions and methods for the detection of hematological
XX PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
XX PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX PS Claim 1; Pages 619-620; 1252pp; English.
XX CC The present invention relates to compositions and methods for the
XX CC detection, diagnosis and therapy of haematological malignancies. The
XX CC present sequence is the protein sequence of a human haematological
XX CC malignancy related antigen. The methods of the present invention comprise
XX CC detecting the presence of haematological malignancy related antigen(s) in
XX CC a sample obtained from the patient (an increased level of the
XX CC polypeptide, compared to an unaffected individual, is indicative of an
XX CC increased risk). Haematological malignancies which can be treated using
XX CC the present invention are chronic lymphocytic leukaemia, lymphoma,
XX CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
XX CC cell non-Hodgkin's lymphoma.
XX SO Sequence 150 AA:
XX
XX Query Match 87.2%; Score 212; DB 22; Length 150;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-20;
XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 KMESLNFI RAHTPYININCEPANSEKNSPSTQCYCS 44
DB 1 KMESLNFI RAHTPYININCEPANSEKNSPSTQCYCS 38

```

```

RESULT 10
AAM80569
ID AAM80569 standard; Protein: 150 AA.
XX AC AAM80569;
XX
XX DT 13-NOV-2001 (first entry)
XX DE Human haematological malignancy-related antigen #308.
XX

```

```

XX Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
KM haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
XX
OS Homo sapiens.
XX
PN WO200164886-A2.
XX
PD 07-SEP-2001.
XX
PF 01-MAR-2001; 2001WO-US07272.
XX
PR 01-MAR-2000; 2000US-0186126.
PR 17-MAR-2000; 2000US-0190479.
PR 27-APR-2000; 2000US-0200545.
PR 28-APR-2000; 2000US-0200303.
PR 28-APR-2000; 2000US-0200779.
PR 01-MAY-2000; 2000US-0200999.
PR 04-MAY-2000; 2000US-0202084.
PR 22-MAY-2000; 2000US-0206201.
PR 14-JUL-2000; 2000US-0218950.
PR 03-AUG-2000; 2000US-0222903.
PR 04-AUG-2000; 2000US-0223416.
PR 07-AUG-2000; 2000US-0223378.
XX
PA (CORI-) CORIXA CORP.
XX
PI Gaiger A, Algate PA, Mannion J;
XX
DR WPI; 2001-514842/56.
XX
PT Compositions and methods for the detection of hematological
PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX
PS Claim 1; Pages 637-638; 1252pp; English.
XX
CC The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the protein sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma.
XX
SQ Sequence 150 AA;
XX
Query Match 87.2%; Score 212; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 KMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 44
DB 1 KMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 38

```

```

KM follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
XX
OS Homo sapiens.
XX
PN WO200164886-A2.
XX
PD 07-SEP-2001.
XX
PF 01-MAR-2001; 2001WO-US07272.
XX
PR 01-MAR-2000; 2000US-0186126.
PR 17-MAR-2000; 2000US-0190479.
PR 27-APR-2000; 2000US-0200545.
PR 28-APR-2000; 2000US-0200303.
PR 28-APR-2000; 2000US-0200779.
PR 01-MAY-2000; 2000US-0200999.
PR 04-MAY-2000; 2000US-0202084.
PR 22-MAY-2000; 2000US-0206201.
PR 14-JUL-2000; 2000US-0218950.
PR 03-AUG-2000; 2000US-0222903.
PR 04-AUG-2000; 2000US-0223416.
PR 07-AUG-2000; 2000US-0223378.
XX
PA (CORI-) CORIXA CORP.
XX
PI Gaiger A, Algate PA, Mannion J;
XX
DR WPI; 2001-514842/56.
XX
PT Compositions and methods for the detection of hematological
PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX
PS Claim 1; Page 642; 1252pp; English.
XX
CC The present invention relates to compositions and methods for the
CC detection, diagnosis and therapy of haematological malignancies. The
CC present sequence is the protein sequence of a human haematological
CC malignancy related antigen. The methods of the present invention comprise
CC detecting the presence of haematological malignancy related antigen(s) in
CC a sample obtained from the patient (an increased level of the
CC polypeptide, compared to an unaffected individual, is indicative of an
CC increased risk). Haematological malignancies which can be treated using
CC the present invention are chronic lymphocytic leukaemia, lymphoma,
CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC cell non-Hodgkin's lymphoma.
XX
SQ Sequence 150 AA;
XX
Query Match 87.2%; Score 212; DB 22; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 KMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 44
DB 1 KMSLNFIKRAHTPYININCEPANPSEKNSPSTQYCS 38

```

```

RESULT 12
AAM80619
ID AAM80619 standard; Protein; 150 AA.
XX
AC AAM80619;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen #317.
XX
KW Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;

```

```

RESULT 13
AAM80950
ID AAM80950 standard; Protein; 150 AA.
XX
AC AAM80950;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human haematological malignancy-related antigen #648.
XX
KW Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
XX
OS Homo sapiens.

```

```
XX WO200164886-A2.
PN
XX
PD 07-SEP-2001.
PE
PF 01-MAR-2001; 2001WO-US07272.
PG
PH
PI
PJ
PK 01-MAR-2000; 2000US-0186126.
PL 17-MAR-2000; 2000US-0190479.
PM 27-APR-2000; 2000US-0200545.
PN 28-APR-2000; 2000US-0200303.
PO 28-APR-2000; 2000US-0200779.
PP 01-MAY-2000; 2000US-0200999.
PQ 04-MAY-2000; 2000US-0202084.
PR 22-MAY-2000; 2000US-0206201.
PS 14-JUL-2000; 2000US-0218950.
PT 03-AUG-2000; 2000US-0222903.
PU 04-AUG-2000; 2000US-0223416.
PV 07-AUG-2000; 2000US-0223378.
PW
PX
PY (CORI-) CORIXA CORP.
PZ
QA
QB Gaiger A, Algate PA, Mannion J;
QC WPI; 2001-514842/56.
QE
QF
QG Compositions and methods for the detection of hematological
QH malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
QJ Lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
QK Claim 1; Pages 773-774; 1252pp; English.
QL
QM The present invention relates to compositions and methods for the
QN detection, diagnosis and therapy of haematological malignancies. The
NQ present sequence is the protein sequence of a human haematological
OQ malignancy related antigen. The methods of the present invention comprise
OR detecting the presence of haematological malignancy related antigen(s) in
OS a sample obtained from the patient (an increased level of the
OT polypeptide, compared to an unaffected individual, is indicative of an
OU increased risk). Haematological malignancies which can be treated using
OV the present invention are chronic lymphocytic leukaemia, lymphoma,
OW follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
OX cell non-Hodgkin's lymphoma.
OY Sequence 150 AA:
PA Best Local Similarity 87.2%; Score 212; DB 22; Length 150;
PB Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
PC
PD Query Match
PE
PF 7 KMESLNFTRAHPTNYINYNCEPANSEKNSPSTGYCYCS 44
PG |||||||
PH Db 1 KMESLNFTRAHPTNYINYNCEPANSEKNSPSTGYCYCS 38
PI
PJ
PK RESULT 14
PL ID AAM81576 standard; Protein: 150 AA.
PM AAM81576:
PN
PO 13-NOV-2001 (first entry)
PQ
PR Human haematological malignancy-related antigen #1274.
PS DE
PT AC
PU AD
PV DT
PW 13-NOV-2001 (first entry)
PX
PY Human: cytostatic; vascular; gene therapy; vaccine; lymphoma;
PZ haematological malignancy; antigen; chronic lymphocytic leukaemia;
QA follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
QB Homo sapiens.
QC OS
QD XN
QE XX
QF MO200164886-A2.
```

PD	07-SEP-2001.
XX	
PF	01-MAR-2001; 2001WO-USO7272.
XX	
PR	01-MAR-2000; 2000US-0186126.
PR	17-MAR-2000; 2000US-0190479.
PR	27-APR-2000; 2000US-0200545.
PR	28-APR-2000; 2000US-0200303.
PR	28-APR-2000; 2000US-0200779.
PR	01-MAY-2000; 2000US-0200999.
PR	04-MAY-2000; 2000US-0202084.
PR	22-MAY-2000; 2000US-0206201.
PR	14-JUL-2000; 2000US-0218950.
PR	03-AUG-2000; 2000US-0223903.
PR	04-AUG-2000; 2000US-0223416.
PR	07-AUG-2000; 2000US-0223378.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PJ	Gaiger A, Algate PA, Mannion J;
PI	
XX	WPI; 2001-514842/56.
DR	
XX	
PT	Compositions and methods for the detection of hematological
PT	malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
PT	lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
XX	
PS	Claim 1; Pages 1019-1020; 1252pp; English.
XX	
CC	The present invention relates to compositions and methods for the
CC	detection, diagnosis and therapy of haematological malignancies. The
CC	present sequence is the protein sequence of a human haematological
CC	malignancy related antigen. The methods of the present invention comprise
CC	detecting the presence of haematological malignancy related antigen(s) in
CC	a sample obtained from the patient (an increased level of the
CC	polypeptide, compared to an unaffected individual, is indicative of an
CC	increased risk). Haematological malignancies which can be treated using
CC	the present invention are chronic lymphocytic leukaemia, lymphoma,
CC	follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
CC	cell non-Hodgkin's lymphoma.
XX	
SO	Sequence 150 AA;
	Query Match 87.2%; Score 212; DB 22; Length 150;
	Best Local Similarity 100.0%; Pred. No. 1.8e-20;
	Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	7 KMESLNFIRAHTPYININCPEANSEKNSPSTQCYCS 44
Dd	1 KMSLNFIKAHTPYINITNCPANSEKNSPSTQICYCS 38
	RESULT 15
ID	AAM80488
AC	AAM80488 standard; Protein: 146 AA.
XX	
AA	AAM80488;
XX	
DT	13-NOV-2001 (first entry)
DE	
XX	Human haematological malignancy-related antigen #186.
XX	
KM	Human; cytotostatic; vascular; gene therapy; vaccine; lymphoma;
KW	haematological malignancy; antigen; chronic lymphocytic leukaemia;
XX	follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma.
OS	Homo sapiens.
XX	
PN	WO200164886-A2.
XX	
PD	07-SEP-2001.
XX	
PF	01-MAR-2001; 2001WO-USO7272.

XX 01-MAR-2000; 2000US-0186126.
 PR 17-MAR-2000; 2000US-0190479.
 PR 27-APR-2000; 2000US-0200545.
 PR 28-APR-2000; 2000US-0200303.
 PR 28-APR-2000; 2000US-0200779.
 PR 01-MAY-2000; 2000US-0200999.
 PR 04-MAY-2000; 2000US-0202084.
 PR 22-MAY-2000; 2000US-0206201.
 PR 14-JUL-2000; 2000US-0218950.
 PR 03-AUG-2000; 2000US-0222903.
 PR 04-AUG-2000; 2000US-0223416.
 PR 07-AUG-2000; 2000US-0223378.

XX (CORI-) CORIXA CORP.
 PA

XX Gaiger A, Algate PA, Mannion J;
 PI

XX WPI; 2001-514842/56.
 DR

XX Compositions and methods for the detection of hematological
 PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular
 PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma -
 XX

PS Claim 1; Pages 584-585; 1252pp; English.
 XX

CC The present invention relates to compositions and methods for the
 CC detection, diagnosis and therapy of hematological malignancies. The
 CC present sequence is the protein sequence of a human hematological
 CC malignancy related antigen. The methods of the present invention comprise
 CC detecting the presence of hematological malignancy related antigen(s) in
 CC a sample obtained from the patient (an increased level of the
 CC polypeptide, compared to an unaffected individual, is indicative of an
 CC increased risk). Hematological malignancies which can be treated using
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B
 CC cell non-Hodgkin's lymphoma.
 XX

XX Sequence 146 AA;
 SQ

5 Query Match 80.7%; Score 196; DB 22; Length 146;
 Best Local Similarity 94.7%; Pred. No. 2,4e-18;

Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 KMESLNFI RAHTPYININCEPANSEKNSPSTOYCS 44
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1 KMESLNFI RAHTPYININCEPANSEKNSPSTOYXS 38

Search completed: March 19, 2003, 17:19:54
 Job time : 45.5 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 19, 2003, 17:22:55 ; Search time 12 Seconds
(without alignments)
107.884 Million cell updates/sec

Title: US-09-674-975-1

Perfect score: 243
Sequence: 1 KISHFLKMSLNFIKRAHPTV.....NCEPAMPSEKNSPTQYCS 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2.6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2.6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2.6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2.6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2.6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2.6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.5	24.9	2391	2 US-08-446-855A-2	Sequence 2, Appl1
2	60.5	24.9	2391	4 US-09-150-741-2	Sequence 2, Appl1
3	56	23.0	101	4 US-09-134-001C-5086	Sequence 5086, Ap
4	54.5	22.4	554	1 US-08-445-586-7	Sequence 7, Appl1
5	54.5	22.4	556	1 US-08-445-586-2	Sequence 2, Appl1
6	53.5	22.0	393	4 US-08-967-024C-25	Sequence 25, Appl1
7	51.5	21.2	551	1 US-08-484-493-15	Sequence 15, Appl1
8	51.5	21.2	551	1 US-08-484-494-15	Sequence 15, Appl1
9	51.5	21.2	551	2 US-08-345-212-15	Sequence 15, Appl1
10	51.5	21.2	551	4 US-09-249-003-15	Sequence 15, Appl1
11	50.5	20.8	149	4 US-09-227-357-443	Sequence 443, App
12	50	20.6	881	4 US-08-960-048-8	Sequence 8, Appl1
13	49.5	20.4	119	1 US-08-469-486-50	Sequence 50, Appl1
14	49.5	20.4	119	2 US-08-469-658-50	Sequence 50, Appl1
15	49.5	20.4	533	4 US-09-360-197-10	Sequence 10, Appl1
16	49.5	20.4	854	4 US-09-268-480-13	Sequence 13, Appl1
17	49	20.2	1261	3 US-09-080-855-2	Sequence 2, Appl1
18	48	19.8	112	4 US-09-134-001C-3131	Sequence 3131, Ap
19	48	19.8	372	3 US-08-918-249-2	Sequence 2, Appl1
20	48	19.8	372	3 US-08-918-249-4	Sequence 4, Appl1
21	48	19.8	372	4 US-09-345-603-2	Sequence 2, Appl1
22	48	19.8	372	4 US-09-345-603-4	Sequence 4, Appl1
23	48	19.8	463	1 US-08-142-439A-2	Sequence 2, Appl1
24	48	19.8	463	2 US-08-869-477-2	Sequence 2, Appl1
25	48	19.8	730	2 US-08-696-944-2	Sequence 2, Appl1
26	48	19.8	772	2 US-08-410-784A-5	Sequence 5, Appl1
27	47.5	19.5	437	1 US-08-136-119-2	Sequence 2, Appl1

28	47.5	19.5	437	2 US-08-481-814A-7	Sequence 7, Appl1
29	47.5	19.5	944	4 US-09-449-285A-2	Sequence 2, Appl1
30	47	19.3	126	1 US-08-170-360-3	Sequence 3, Appl1
31	47	19.3	126	2 US-08-888-497-38	Sequence 38, Appl1
32	47	19.3	126	4 US-09-362-230-38	Sequence 38, Appl1
33	47	19.3	126	5 PCT-US94-07926-38	Sequence 38, Appl1
34	47	19.3	148	2 US-08-888-497-36	Sequence 36, Appl1
35	47	19.3	148	4 US-09-362-230-36	Sequence 36, Appl1
36	47	19.3	148	5 PCT-US94-07926-36	Sequence 36, Appl1
37	47	19.3	222	1 US-08-126-593A-2	Sequence 2, Appl1
38	47	19.3	222	1 US-08-454-039A-2	Sequence 2, Appl1
39	47	19.3	349	4 US-08-462-467B-18	Sequence 18, Appl1
40	47	19.3	554	4 US-08-462-467B-22	Sequence 22, Appl1
41	47	19.3	730	1 US-08-430-235B-17	Sequence 17, Appl1
42	47	19.3	730	4 US-08-793-624-17	Sequence 17, Appl1
43	47	19.3	730	5 PCT-US95-10194-17	Sequence 17, Appl1
44	47	19.3	830	4 US-09-562-737-39	Sequence 39, Appl1
45	47	19.3	869	2 US-08-483-101-15	Sequence 15, Appl1

ALIGNMENTS

RESULT 1
US-08-446-855A-2
Sequence 2, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
TITLE OF INVENTION: Nucleotide sequence encoding carbanoyl
phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
City: Arlington
State: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2391 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-855A-2
Query Match 24.9%; Score 60.5; DB 2; Length 2391;
Best Local Similarity 34.3%; Pred. No. 12;
Matches 12; Conservative 10; Mismatches 12; Indels 1; Gaps 1;
OY 2 ISHFLKMSLNFIKRAHPTV-NIYNCEPAMPSEKN 35
Db 317 INHTLLRDKMNLITSESEYELKDLHNCNFSNDDKN 351

```

RESULT 2
US-09-150-741-2
; Sequence 2, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Catramoyl Phosphate
; Patent No. 6183996
; TITLE OF INVENTION: Synthetase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2

Query Match          24.3%; Score 60.5; DB 4; Length 2391;
Best Local Similarity 34.3%; Pred. NO. 12;
Matches 12; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

Oy      2 ISHFLKMSLNFIIRHTPYI-NIYNCEPANPSEKN 35
       | : | | : | | : | : ||| : | | |
Db    317 IHNLTLLRDKMNLITSEXYLKDLHCNCFNSSDKN 351


RESULT 3
US-09-134-001C-5086
; Sequence 5086, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5086
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5086

Query Match          23.0%; Score 56; DB 4; Length 101;
Best Local Similarity 35.7%; Pred. NO. 11;
Matches 15; Conservative 5; Mismatches 18; Indels 4; Gaps 2;

Oy      6 LKMESLN--FIRAHPYT-NIYNCEPANPSEKSPSTQCY 43
       | | : | : | | : | | | | | | | |
Db      9 LKYNEIKNNOLISSETPLENFQDCIKKKNKISOSTLCNY 50


RESULT 4
US-08-445-586-7
; Sequence 7, Application US/08445586

```

```

: Patent No.5627050
: GENERAL INFORMATION:
: APPLICANT: Takeshita, Sunao
: APPLICANT: Ito, Toshimi
: APPLICANT: Otawara-Hamamoto, Yoko
: APPLICANT: Amann, Egon
: TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
: TITLE OF INVENTION: Process for its Production
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESS: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/445,586
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/111,887
: FILING DATE: 26-AUG-1993
: APPLICATION NUMBER: JP 230030/92
: FILING DATE: 28-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 324034/92
: FILING DATE: 03-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Forman, David S.
: REGISTRATION NUMBER: 33,694
: REFERENCE/DOCKET NUMBER: 02481.1322-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 554 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-445-586-7

Query Match 22.4%; Score 54.5; DB 1; Length 554;
Best Local Similarity 40.7%; Pred. No.14;
Matches 11; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

QY 3 SHFLMESLNFIRAH-TPYININYCEP 28
I : : : : : : : : : : : : : : :
Db 55 SQVXKLPFINFMRAGHTFLNAYTNSP 81

RESULT 5
US-08-445-586-2
: Sequence 2, Application US/08445586
: Patent No. 5627050
: GENERAL INFORMATION:
: APPLICANT: Takeshita, Sunao
: APPLICANT: Ito, Toshimi
: APPLICANT: Otawara-Hamamoto, Yoko
: APPLICANT: Amann, Egon
: TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
: TITLE OF INVENTION: Process for its Production
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESS: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/445,586
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/111,887
: FILING DATE: 26-AUG-1993
: APPLICATION NUMBER: JP 230030/92
: FILING DATE: 28-AUG-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 324034/92
: FILING DATE: 03-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Forman, David S.
: REGISTRATION NUMBER: 33,694
: REFERENCE/DOCKET NUMBER: 02481.1322-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 554 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-445-586-7

```



```

APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 22.0%; Score 53.5; DB 4; Length 393;
Best Local Similarity 29.4%; Pred. No. 13;
Matches 15; Conservative 7; Mismatches 14; Indels 15; Gaps 2

OY 1 KISHLKKESLFTRAHPIYINICEPA-----NPEKKSPSTQY 41
Db 346 RVSHLP-----WIRSHTKENGALSPVAPRPFLINPDKYEPPEEY 391
:::|||||:::|||||:::|
:::|||||:::|||||:::|
:::|||||:::|||||:::|

RESULT 7
US-08-484-493-15
Sequence 15, Application US/08484493
Patent No. 5728381
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Biellicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
NUMBER OF INVENTION: IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,493
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-493-15

Query Match 21.2%; Score 51.5; DB 1; Length 551;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 10; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 3 SHFLKMSLNFIRAHPTPIYINCEPANPSEKNSPSTQYCY 43
DB 170 AHLPEFHGDFVGHNLPTNSMCDDTG-LHKDPDSQRCY 209

RESULT 8
US-08-484-494-15
Sequence 15, Application US/08484494
Patent No. 5798239
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Biellikl, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4366
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-494-15

Query Match 21.2%; Score 51.5; DB 1; Length 551;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 10; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 3 SHFLKMSLNFIRAHPTPIYINCEPANPSEKNSPSTQYCY 43
DB 170 AHLPEFHGDFVGHNLPTNSMCDDTG-LHKDPDSQRCY 209

RESULT 9
US-08-345-212-15
Sequence 15, Application US/08345212
Patent No. 5932211
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Biellikl, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
NUMBER OF SEQUENCES: 15
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4366
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-345-212-15

Query Match 21.2%; Score 51.5; DB 2; Length 551;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 10; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 3 SHFLKMSLNFIRAHPTPIYINCEPANPSEKNSPSTQYCY 43
DB 170 AHLPEFHGDFVGHNLPTNSMCDDTG-LHKDPDSQRCY 209

RESULT 10
US-09-249-003-15
Sequence 15, Application US/09249003
Patent No. 6153188
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Biellikl, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J

```
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; TITLE OF INVENTION: IDURONATE 2-SULFATASE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,003
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-249-003-15

Query Match 21.2%; Score 51.5; DB 4; Length 551;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 10; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 3 SHFLKMSLNFIRAHPTVINYNCPEANPSEKNSPTQYCY 43
Db 170 AHPFNHGFVGHNLFTSMSCDDTG-LHKPFPSQRCY 209

RESULT 11
; US-09-227-357-443
; Sequence 443, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
```

```
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 443
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-227-357-443

Query Match 20.8%; Score 50.5; DB 4; Length 149;
Best Local Similarity 35.1%; Pred. No. 11;
Matches 13; Conservative 5; Mismatches 8; Indels 11; Gaps 2;

QY 15 RAHTPYINIYNC-----EPANPSEKNSPTQYCY 42
Db 61 RANNPISNVSSCPIDGPFYADPNYASEK--PSTAFYC 95

RESULT 12
; US-08-960-048-8
; Sequence 8, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
```

APPLICANT: Stalker, D. et al.
TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
FILE REFERENCE: 15621/01/US
CURRENT APPLICATION NUMBER: US/08/960,048C
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/029,987
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 8
LENGTH: 881
TYPE: PRT
ORGANISM: Oryzae sativa
US-08-960-048-8

Query Match 20.6%; Score 50; DB 4; Length 881;
Best Local Similarity 33.3%; Pred. No. 1.1e+02;
Matches 11; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

OY 9 ESNLRHTYININCEPANPSEKNSPSTOY 41
DB 208 ETSEFARKVPFKYNIETPRAPMYFSQKIDY 240

RESULT 13
US-08-469-486-50
Sequence 50, Application US/08469486
Patent No. 5739281
GENERAL INFORMATION:
APPLICANT: Thøgersen, Hans Christian
APPLICANT: Holte, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: Improved method for the refolding of
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,486
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-486-50

Query Match 20.4%; Score 49.5; DB 1; Length 119;
Best Local Similarity 34.1%; Pred. No. 11;
Matches 14; Conservative 6; Mismatches 16; Indels 5; Gaps 2;

OY 5 FLKMSLNFIRA--HTPYININCEPANPSEKNSPSTOYCY 43
DB 9 FLVLVSLTGLYAIQKTPQIOYVS---RHPPENCKPILNCY 46

RESULT 14
US-08-469-658-50
Sequence 50, Application US/08469658
Patent No. 5917018
GENERAL INFORMATION:
APPLICANT: Thøgersen, Hans Christian
APPLICANT: Holte, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,658
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-658-50

Query Match 20.4%; Score 49.5; DB 2; Length 119;
Best Local Similarity 34.1%; Pred. No. 11;
Matches 14; Conservative 6; Mismatches 16; Indels 5; Gaps 2;

OY 5 FLKMSLNFIRA--HTPYININCEPANPSEKNSPSTOYCY 43
DB 9 FLVLVSLTGLYAIQKTPQIOYVS---RHPPENCKPILNCY 46

RESULT 15
US-09-360-197-10
Sequence 10, Application US/09360197
Patent No. 6287859
GENERAL INFORMATION:
APPLICANT: Bassilana, Frederic

; APPLICANT: Iazdunski, Michel
; APPLICANT: Waldmann, Rainer
; APPLICANT: Deweille, Jan R.
; TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive
; TITLE OF INVENTION: Cationic Channels, Their Cloning and Applications
; FILE REFERENCE: 989,6706P
; CURRENT APPLICATION NUMBER: US/09/360,197
; CURRENT FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: 09/129,758
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/095,408
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 533
; TYPE: PRT
; ORGANISM: rattus sp.
US-09-360-197-10

Query Match 20.4%; Score 49.5; DB 4; Length 533;
Best Local Similarity 31.7%; Pred. No. 69;

Matches 13; Conservative 7; Mismatches 18; Indels 3; Gaps 1;

OY 6 LKMSLNFIRAHPTPIYINICEPAMP--SEKNSPSTQYCY 43

DB 488 LLOELNGHRTVPHLSLGPPRPPTPCAVTKTLSASHRTCY 528

Search completed: March 19, 2003, 17:23:31
Job time : 14 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2003, 17:30:39 ; Search time 17 Seconds
(without alignments)
138.373 Million cell updates/sec

Title: US-09-674-975-1

Perfect score: 243
Sequence: 1 KISHFLKMSLNFRAHPTPY.....NCEPANPSEKSPSTQYCS 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
2	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
3	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
4	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
5	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
6	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
7	212	87.2	150	9	US-09-796-692-928 Sequence 928, App
8	184	75.7	94	9	US-09-796-692-928 Sequence 928, App
9	144	59.3	137	9	US-09-796-692-928 Sequence 928, App
10	97	39.9	141	9	US-09-796-692-928 Sequence 928, App
11	61.5	25.3	1013	9	US-10-121-049-38 Sequence 38, App1
12	61.5	25.3	1013	9	US-10-121-049-38 Sequence 38, App1
13	61.5	25.3	1013	9	US-10-123-904-38 Sequence 38, App1
14	61.5	25.3	1013	9	US-10-140-470-38 Sequence 38, App1
15	61.5	25.3	1013	9	US-10-175-746-38 Sequence 38, App1
16	61.5	25.3	1013	9	US-10-175-746-38 Sequence 38, App1
17	61.5	25.3	1013	9	US-10-176-921-38 Sequence 38, App1
18	61.5	25.3	1013	9	US-10-137-865-38 Sequence 38, App1
19	61.5	25.3	1013	9	US-10-140-474-38 Sequence 38, App1

20	61.5	25.3	1013	9	US-10-142-431-38 Sequence 38, App1
21	61.5	25.3	1013	9	US-10-143-114-38 Sequence 38, App1
22	61.5	25.3	1013	9	US-10-140-002-38 Sequence 38, App1
23	61.5	25.3	1013	9	US-10-142-419-38 Sequence 38, App1
24	61.5	25.3	1013	9	US-10-123-262-38 Sequence 38, App1
25	61.5	25.3	1013	9	US-10-142-423-38 Sequence 38, App1
26	52.5	21.6	177	9	US-09-791-279-128 Sequence 128, App
27	52	21.4	532	10	US-09-801-368-360 Sequence 360, App
28	50.5	20.8	149	9	US-09-983-802-443 Sequence 443, App
29	50.5	20.8	453	10	US-09-815-242-5668 Sequence 5668, App
30	50.5	20.8	453	10	US-09-815-242-12701 Sequence 12701, App
31	50	20.6	881	10	US-09-838-539-8 Sequence 8, App1
32	49.5	20.4	117	10	US-09-799-514-10 Sequence 10, App1
33	49.5	20.4	533	10	US-09-864-761-37151 Sequence 37151, App
34	49	20.2	271	10	US-09-745-761-201 Sequence 201, App
35	49	20.2	305	10	US-09-815-242-12482 Sequence 12482, App
36	49	20.2	322	10	US-09-815-242-5327 Sequence 5327, App
37	49	20.2	468	10	US-09-801-368-398 Sequence 398, App
38	48	19.8	64	10	US-09-925-300-1765 Sequence 1765, App
39	48	19.8	91	10	US-09-925-297-750 Sequence 750, App
40	48	19.8	161	10	US-09-764-877-1289 Sequence 1289, App
41	47.5	19.5	236	10	US-09-923-300-1410 Sequence 1410, App
42	47.5	19.5	944	10	US-09-964-238-2 Sequence 2, App1
43	47	19.3	378	10	US-09-801-368-330 Sequence 330, App
44	47	19.3	980	9	US-10-108-605-195 Sequence 195, App
45	46.5	19.1	489	10	US-09-748-127-4 Sequence 4, App1

ALIGNMENTS

RESULT 1

US-09-796-692-928

Sequence 928, Application US/09796692

Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Mannon, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

FILE REFERENCE: 2077, 001200

CURRENT APPLICATION NUMBER: US/09/796,692

CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 928

LENGTH: 150

TYPE: PRT

ORGANISM: Homo sapiens
US-09-796-692-928

Query Match 87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 44
|||||
DB 1 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 38

RESULT 2
US-09-796-692-933

Sequence 933, Application US/09796692
Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 933

LENGTH: 150

TYPE: PRT

ORGANISM: Homo sapiens

US-09-796-692-933

Query Match 87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 44
|||||
DB 1 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 38

RESULT 3

US-09-796-692-974

Sequence 974, Application US/09796692

Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/223,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,378

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 974

LENGTH: 150

TYPE: PRT

ORGANISM: Homo sapiens

US-09-796-692-974

Query Match 87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 44
|||||
DB 1 KMESLNFIRAHPTPIINYNCBPANPSEKNSPSTOYCYS 38

RESULT 4

US-09-796-692-983

Sequence 983, Application US/09796692

Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692

PRIOR FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084


```

; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 983
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(150)
; OTHER INFORMATION: Xaa = Any amino acid
; US-09-796-692-983
```

```

Query Match      87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      7 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 44
Db      1 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 38
```

```

RESULT 5
US-09-796-692-1314
```

```

; Sequence 1314, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1314
; LENGTH: 150
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-692-1314
```

```

Query Match      87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      7 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 44
Db      1 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 38
```

```

RESULT 6
US-09-796-692-1940
```

```

; Sequence 1940, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1940
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(150)
; OTHER INFORMATION: Xaa = Any amino acid
; US-09-796-692-1940
```

```

Query Match      87.2%; Score 212; DB 9; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      7 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 44
Db      1 KMSLNFIRAHPTPIINYNCPEANPSEKNSPTQYCS 38
```

```

RESULT 7
US-09-796-692-852
```

```
Sequence 852, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 852
LENGTH: 146
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(146)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-852

Query Match      80.7%; Score 196; DB 9; Length 146;
Best Local Similarity 94.7%; Pred. No. 3,1e-18;
Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 KMSLNFRATPYININCEPANPSEKNSPTQYCS 44
Db 1 KMSLNFRATPYININCEPANPSEKNSPTQYXS 38

RESULT 8
US-09-796-692-1260
Sequence 1260, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
```

```
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1260
LENGTH: 94
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(94)
OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-1260

Query Match      75.7%; Score 184; DB 9; Length 94;
Best Local Similarity 89.5%; Pred. No. 6.8e-17;
Matches 34; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 KMSLNFRATPYININCEPANPSEKNSPTQYCS 44
Db 1 KMSLNFRATPYININCEPANPSEKNSPTQYXS 38

RESULT 9
US-09-796-692-1320
Sequence 1320, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
```

;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1320
;; LENGTH: 137
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-796-692-1320

Query Match 59.3%; Score 144; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 YINYNCEPANPSEKNSPSTQCYCS 44
Db 1 YINYNCEPANPSEKNSPSTQCYCS 25

RESULT 10
US-09-796-692-820
;; Sequence 820, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; FILE REFERENCE: 2077,001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 820
;; LENGTH: 141
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: variant
;; LOCATION: (1)...(141)
;; OTHER INFORMATION: Xaa = Any amino acid
US-09-796-692-820

Query Match 39.9%; Score 97; DB 9; Length 141;
Best Local Similarity 94.4%; Pred. No. 2.1e-05;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 27 EPANPSEKNSPSTQCYCS 44
Db 1 EPANPSEKNSPSTQCYCS 18

RESULT 11
US-10-028-072-38
;; Sequence 38, Application US/10028072
;; Publication No. US20030004311A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: Deforge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Flivaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gunney, Austin L.
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang
;; TITLE OF INVENTION:
;; FILE REFERENCE:
;; CURRENT APPLICATION NUMBER: US/10/028,072
;; CURRENT FILING DATE: 2001-12-19
;; PRIOR APPLICATION NUMBER: 60/049911
;; PRIOR FILING DATE: 1997-06-18
;; PRIOR APPLICATION NUMBER: 60/056974
;; PRIOR FILING DATE: 1997-08-26
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059115
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059117
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059122
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059184
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059263
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; PRIOR APPLICATION NUMBER: 60/059588
;; PRIOR FILING DATE: 1997-09-19
;; PRIOR APPLICATION NUMBER: 60/059836
;; PRIOR FILING DATE: 1997-09-24
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/062285
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/062814
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/062816
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063045
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063082
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/063127
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063327
;; PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match

Best Local Similarity 25.3%; Score 61.5; DB 9; Length 1013;
Matches 11; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 14 IRAHTPYININCEPANPSEKNSPSTQYCS 44
Db 627 LKAHQY-GVOACVPCGPGTKNKHISLCYN 656

```
RESULT 12
US-10-121-049-38
; Sequence 38, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 38
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 877, 882
; OTHER INFORMATION: unknown amino acid
US-10-121-049-38

; Query Match
; Best Local Similarity 25.3%; Score 61.5; DB 9; Length 1013;
; Matches 11; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 14 IRAHTPYININCEPANPSEKNSPSTQYCS 44
Db 627 LKAHQY-GVQACVPCGPGTKNKKIHSLCYN 656

RESULT 13
US-10-123-904-38
; Sequence 38, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
```

```
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 38
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 877, 882
; OTHER INFORMATION: unknown amino acid
US-10-123-904-38

; Query Match
; Best Local Similarity 25.3%; Score 61.5; DB 9; Length 1013;
; Matches 11; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 14 IRAHTPYININCEPANPSEKNSPSTQYCS 44
Db 627 LKAHQY-GVQACVPCGPGTKNKKIHSLCYN 656

RESULT 14
US-10-140-470-38
; Sequence 38, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 38
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 877, 882
; OTHER INFORMATION: unknown amino acid
US-10-140-470-38

; Query Match
; Best Local Similarity 25.3%; Score 61.5; DB 9; Length 1013;
; Matches 11; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 14 IRAHTPYININCEPANPSEKNSPSTQYCS 44
Db 627 LKAHQY-GVQACVPCGPGTKNKKIHSLCYN 656

RESULT 15
US-10-175-746-38
```

```
; Sequence 38, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mel-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 38
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 877..882
; OTHER INFORMATION: unknown amino acid
US-10-175-746-38

Query Match 25.3%; Score 61.5; DB 9; Length 1013;
Best Local Similarity 35.5%; Pred. No. 7.7;
Matches 11; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 14 IRAHTPYINIVCEPANPSEKNSPSTQYCYX 44
DB 627 LKAHQPY-GVQACVPCGPGTKNNKIHSLCYN 656
```

Search completed: March 19, 2003, 17:57:29
Job time : 27.5 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2003, 17:22:00 ; Search time 13.5 seconds
(without alignments)
313.327 Million cell updates/sec

Title: US-09-674-975-1

Perfect score: 243
Sequence: 1 KISHFLKMSLNFIHAPY.....NCEPANPSEKNSPTQCYCS 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have
score greater than or equal to the score of the result being
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Desc
1	243	100.0	297	1 A30586	B-C
2	160	65.8	291	2 A30558	B-C
3	61	25.1	702	2 D86469	B-C
4	60.5	24.9	2391	2 T18410	B-C
5	57.5	23.7	443	2 S44786	B-C
6	56.5	23.3	1643	1 RRMGNV	B-C
7	56.5	23.3	2425	2 D69426	B-C
8	56	23.0	711	2 E85098	B-C
9	55.5	22.8	112	2 S35880	B-C
10	55.5	22.8	2150	2 T32497	B-C
11	54.5	22.4	547	2 T00977	B-C
12	54	22.2	253	2 F69134	B-C
13	54	22.2	329	2 H71192	B-C
14	54	22.2	337	2 T23892	B-C
15	54	22.2	447	2 B81519	B-C
16	54	22.2	447	2 F72021	B-C
17	54	22.2	447	2 A86603	B-C
18	54	22.2	2091	2 A97077	B-C
19	53	21.8	332	2 C95150	B-C
20	53	21.8	4981	2 T18489	B-C
21	52.5	21.6	117	2 S67095	B-C
22	52.5	21.6	331	2 T22873	B-C
23	52.5	21.6	604	2 S39885	B-C
24	52	21.4	411	2 S58105	B-C
25	52	21.4	532	2 JC6170	B-C
26	52	21.4	650	2 T46660	B-C
27	52	21.4	1576	2 T21172	B-C
28	51.5	21.2	438	2 G96578	B-C
29	51.5	21.2	551	2 S01793	B-C

30	51	21.0	197	2 C97830	RP681 homolog [imp
31	51	21.0	355	2 E81415	DNA-directed DNA p
32	51	21.0	427	1 ZFBG4	gene F protein - p
33	51	21.0	434	2 T37125	hypothetical prote
34	51	21.0	462	2 C64500	hypothetical prote
35	51	21.0	513	2 S50915	SIN3 protein-bindi
36	51	21.0	772	2 E96977	hypothetical prote
37	51	21.0	2109	1 I50421	aggreacan precursor
38	50.5	20.8	288	2 S46536	chitinase (EC 3.2.
39	50.5	20.8	355	2 D97339	dioxygenase relate
40	50.5	20.8	453	2 JC5607	replication initia
41	50.5	20.8	551	2 S07089	arylsulfatase (EC
42	50.5	20.8	744	2 E86255	hypothetical prote
43	50	20.6	213	2 C81035	hypothetical prote
44	50	20.6	227	2 T21298	hypothetical prote
45	50	20.6	431	2 S22330	gene F protein - p

ALIGNMENTS

RESULT 1

A30586
B-cell surface antigen CD20 - human
N:Alternate names: B-lymphocyte antigen CD20; BI
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1989 #sequence: revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: A30586; J10042; A27400; S00387
R:Redder, T.F.; Klejman, G.; Schlossman, S.F.; Salto, H.
J. Immunol. 142, 2560-2568, 1989
A:Title: Structure of the gene encoding the human B lymphocyte differentiation antigen
A:Reference number: A30586; MUID:89176281; PMID:2466899
A:Accession: A30586
A:Molecule type: DNA
A:Residues: 1-297 <TMD>
A:Note: the authors translated the codon ATG for residue 148 as His, TCA for residue
, and GAC for residue 294 as Ser
R:Stamenkovic, I.; Seed, B.
J. Exp. Med. 167, 1975-1980, 1988
A:Title: Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (BI, Bp35)
A:Reference number: J10042; MUID:8828386; PMID:3260267
A:Accession: J10042
A:Molecule type: mRNA
A:Residues: 1-297 <STA>
A:Cross-references: GB:X12530; NID:g29773; PIDN:CAA31046.1; PID:g29774
R:Redder, T.F.; Streuli, M.; Schlossman, S.F.; Salto, H.
Proc. Natl. Acad. Sci. U.S.A. 85, 208-212, 1988
A:Title: Isolation and structure of a cDNA encoding the BI (CD20) cell-surface antigen
A:Reference number: A27400; MUID:88124792; PMID:2448768
A:Accession: A27400
A:Molecule type: mRNA
A:Residues: 1-297 <ME2>
A:Cross-references: GB:M27394; GB:J03574; NID:g179307; PIDN:AAA35581.1; PID:g179308
R:Einfield, D.A.; Brown, J.P.; Valentine, M.A.; Clark, E.A.; Ledbetter, J.A.
EMBO J. 7, 711-717, 1988
A:Title: Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic p
A:Reference number: S00387; MUID:88283639; PMID:2456210
A:Accession: S00387
A:Molecule type: mRNA
A:Residues: 1-12, 14-297 <EIN>
A:Cross-references: EMBL:X07203; NID:g29775; PIDN:CAA30179.1; PID:g29776
A:Experimental source: Daudi cells
C:Comment: This protein appears not to be glycosylated. Isoforms of varying molecular
C:Comment: This protein plays an important role in B cell activation.
C:Genetics:
A:Gene: GDB:CD20
A:Cross-references: GDB:119761; OMIM:112210
A:Map position: 11q12-11q13.1
C:Superfamily: B-cell surface antigen CD20
C:Keywords: B-cell; phosphoprotein; transmembrane protein
F:1-51/Domain: Intracellular #status predicted <CMT1>
F:52-103/Domain: Transmembrane #status predicted <TMD2>
F:104-116/Domain: Intracellular #status predicted <CMT2>

F:117-141/Domain: transmembrane #status predicted <TM>
F:142-185/Domain: extracellular #status predicted <EXT>
F:186-212/Domain: transmembrane #status predicted <TM>
F:213-297/Domain: intracellular #status predicted <CYT>
F:167-189/Disulfide bonds: #status predicted

Query Match 100.0%; Score 243; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 1.8e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISHFLKMSLNPIRAHTPIYINYNCEPANPSEKNSPSTGYCS 44
DB 142 KISHFLKMSLNPIRAHTPIYINYNCEPANPSEKNSPSTGYCS 185

RESULT 2

A:30558
B-cell surface antigen CD20 homolog - mouse
N:Alternate names: B-cell differentiation antigen Ly-44
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 07-Feb-1997
C:Accession: A30558
R:Pedder, T.F.; Kiehlman, G.; Distech, C.M.; Adler, D.A.; Schlossman, S.F.; Salto, H.
J. Immunol. 141, 4388-4394, 1988
A:Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation
A:Reference number: A30558; MUID:89067519; PMID:2461992
A:Accession: A30558
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-291 <TED>
C:Genetics:
A:Map position: 19
C:Superfamily: B-cell surface antigen CD20
C:Keywords: membrane protein; phosphoprotein; surface antigen

Query Match 65.8%; Score 160; DB 2; Length 291;
Best Local Similarity 65.1%; Pred. No. 7.2e-13;
Matches 28; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISHFLKMSLNPIRAHTPIYINYNCEPANPSEKNSPSTGYCS 44
DB 137 LSHFLKMRLELIOTSKPYVDIDCPNSSEKNSPSTGYCS 179

RESULT 3

A:DB6469
protein F12K21.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: DB6469
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chlin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Malt, R.; Matzali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: DB6469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-702 <STO>
A:Cross-references: GB:AE005172; NID:g8778248; PIDN:AAF79257.1; GSPDB:GN00141
C:Genetics:
A:Gene: F12K21.12
A:Map position: 1

Query Match 25.1%; Score 61; DB 2; Length 702;
Best Local Similarity 35.5%; Pred. No. 8.2; Indels 0; Gaps 0;
Matches 11; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 14 IRAHTPIYINYNCEPANPSEKNSPSTGYCS 44
DB 624 VRAHTPMNSLFCMLMENEVEERTPRDQLSFA 654

RESULT 4

A:18410
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) II - malaria parasit
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18410
R:Flores, M.V.C.; O'Sullivan, W.J.; Stewart, T.S.
submitted to the EMBL Data Library, March 1997
A:Description: Characterisation of the carbamoyl phosphate synthetase gene from Plasm
A:Reference number: Z18931
A:Accession: T18410
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2391 <FIO>
A:Cross-references: EMBL:L32150; NID:q476023; PID:q476024; PIDN:AAA29522.1
C:Superfamily: rudimentary enzyme; aspartate/orithine carbamoyltransferase homology;
arabamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homology; carbamoyl-p
F:9672367/Domain: carbamoyl-phosphate synthase (ammonia) homology <CPA>

Query Match 24.9%; Score 60.5; DB 2; Length 2391;
Best Local Similarity 34.3%; Pred. No. 34;
Matches 12; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

QY 2 ISHFLKMSLNPIRAHTPIYINYNCEPANPSEKN 35
DB 317 INHTLRDKMNLITSEEYKDLHNCNFSNSDKN 351

RESULT 5

A:544786
D2007.5 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 28-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C:Accession: 544786
R:Favella, A.D.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the C. elegans cosmid D2007.
A:Reference number: 544786
A:Accession: 544786
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <FAV>
A:Cross-references: EMBL:L16560; NID:q289666; PID:q289667
C:Genetics:
A:Intons: 25/3; 197/2; 250/3; 321/1; 373/3; 419/3
C:Superfamily: Caenorhabditis elegans D2007.5 protein

Query Match 23.7%; Score 57.5; DB 2; Length 443;
Best Local Similarity 45.7%; Pred. No. 14;
Matches 16; Conservative 7; Mismatches 5; Indels 7; Gaps 3;

QY 10 SLNPIRAHT-PIYINYNCEPANPSEKNSPST 39
DB 286 SLADIRNHSFPVNLQSNYN--PANGTKKNSST 318

RESULT 6

A:RRGMV
genome polypeptide - narciissus mosaic virus
N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: narciissus mosaic virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C:Accession: J00470
R:Zuidema, D.; Linthorst, H.J.M.; Huisman, M.J.; Asjes, C.J.; Bol, J.F.
J. Gen. Virol. 70, 267-276, 1989

A>Title: Nucleotide sequence of narciassus mosaic virus RNA.
 A:Reference number: JF0470; MUID:89279206; PMID:2732689
 A:Accession: JF0470
 A:Molecule type: genomic RNA
 A:Residues: 1-1643 <201>
 A:Cross-references: GB:D13747; GB:D00405; NID:g222107; PIDN:BA02891.1; PID:g222108
 C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase
 C:Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; F
 F:868-875/Region: nucleotide-binding motif A (P-loop)
 F:931-936/Region: nucleotide-binding motif B
 F:874/Binding site: ATP (Lys) #status predicted

Query Match 23.3%; Score 56.5; DB 1; Length 1643;
 Best Local Similarity 26.9%; Pred. No. 75;
 Matches 14; Conservative 11; Mismatches 6; Indels 21; Gaps 3;

OY 4 HFLMESINFRHTPTI-----NINCEPAN--PSKNSPSTGY 41
 Db 1107 YMAKLES-----TPYLKAFIDTYRDEKTEVNSOPASAPTEPEAPATHF 1151

RESULT 7
 D69426
 Surface layer protein B (slgB-2) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C:Accession: D69426
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9399475
 A:Accession: D69426
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2425 <KLE>
 A:Cross-references: GB:AE001006; GB:AE000782; NID:g2689329; PIDN:AB89834.1; PID:g264915

Query Match 23.3%; Score 56.5; DB 2; Length 2425;
 Best Local Similarity 30.0%; Pred. No. 1.1e+02;
 Matches 15; Conservative 9; Mismatches 19; Indels 7; Gaps 2;

OY 1 KISHFLMESINFRHTPTIYNIVN---CEPANPSE---KNSPSTGYC 43
 Db 1764 KISSFLARARIEFYRSTPEVEYISVNDLIEVLNOAETALRNPGMPDFY 1813

RESULT 8
 E85098
 hypothetical protein AT4g09630 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: E85098
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprit
 Nature 402, 769-777, 1999
 A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488; PMID:10617198
 A:Accession: E85098
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-711 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267658; PIDN:CAB78086.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4g09630
 A:Map position: 4

Query Match 23.0%; Score 56; DB 2; Length 711;
 Best Local Similarity 32.3%; Pred. No. 36;
 Matches 10; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 14 IRAHTPYININCEPANPSEKNSPSTGYC 44
 Db 632 VREHTPMSNLFSCIMFNEVERFTPRDLSFA 662

RESULT 9
 S35880
 AR0 protein - Indian cassava mosaic virus
 C:Species: Indian cassava mosaic virus
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
 C:Accession: S35880
 R:Hong, Y.
 submitted to the EMBL data library, July 1993
 A:Reference number: S35878
 A:Accession: S35880
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-112 <HON>
 A:Cross-references: EMBL:Z24758; NID:g395351; PIDN:CA80884.1; PID:g395352
 C:Superfamily: tomato yellow leaf curl virus V2 protein

Query Match 22.8%; Score 55.5; DB 2; Length 112;
 Best Local Similarity 33.3%; Pred. No. 6;
 Matches 13; Conservative 8; Mismatches 17; Indels 1; Gaps 1;

OY 5 FLKMSINFRHTPTIYNIVN-CEPANSEKNSPSTGYC 42
 Db 46 FSVIRAKYVATSRHNYFSRLGSSPSSELRDPIQPC 84

RESULT 10
 T32497
 hypothetical protein C08G9.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32497
 R:Geisel, C.; Stellyes, L.
 submitted to the EMBL data library, December 1997
 A:Description: The sequence of C. elegans cosmid C08G9.
 A:Reference number: Z21179
 A:Accession: T32497
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2150 <GEI>
 A:Cross-references: EMBL:AF036687; PIDN:AB88311.1; GSPDB:GN00022; CESP:C08G9.2
 A:Experimental source: strain Bristol N2; clone C08G9
 C:Genetics:
 A:Gene: CESP:C08G9.2
 A:Map position: 4
 A:Introns: 242/1; 306/1; 340/3; 485/1; 545/1; 736/2; 791/1; 829/1; 886/3; 942/1; 1079

Query Match 22.8%; Score 55.5; DB 2; Length 2150;
 Best Local Similarity 41.2%; Pred. No. 1.3e+02;
 Matches 14; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

OY 11 LNFIRHTPTI-NINCEPANPSEKNSPSTGYC 43
 Db 610 LNFCPAGPPTISIGNNSCKDECPRTTWCH 643

RESULT 11
 T00977
 probable pectinesterase At2g26440 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T00977; E84660
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
 submitted to the EMBL data library, April 1998
 A:Description: Arabidopsis thaliana chromosome II BAC T9J22 genomic sequence.
 A:Reference number: Z14161
 A:Accession: T00977
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA

· Thu Mar 20 10:59:51 2003

us-09-674-975-1.rpr

Page 5

```

QY      3 SHELFMESINFRATHP-YININCEPANPSEK-----NSPSTQYCY 43
      1 : 1 : : : : : : : : : : : : : : : : : : : : : : : :
Db      209 STYIAKSKLQLLPKDSPLYLLDRSCSPT--AEKFSFSFNPLADFCY 253

```

Search completed: March 19, 2003, 17:22:42
Job time : 17.5 secs

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X12530; AAA31046.1; -
 DR EMBL: X27394; AAA35581.1; -
 DR EMBL: M27395; -; NOT_ANNOTATED_CDS.
 DR EMBL: L23419; AAA88911.1; -
 DR EMBL: L23415; AAA88911.1; JOINED.
 DR EMBL: L23416; AAA88911.1; JOINED.
 DR EMBL: X07203; AAA30179.1; -
 DR EMBL: X07204; AAA30180.1; -
 DR EMBL: BC002807; AAH02807.1; -
 DR PIR: A27400; A27400.
 DR PIR: J10042; J10042.
 DR PIR: A30586; A30586.
 DR PIR: S00387; S00387.
 DR Genew: HGNC:7315; MS4A1.
 DR MIM: 112210; -
 KW B-cell; Transmembrane; Phosphorylation.
 FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT DOMAIN 210 297 POTENTIAL.
 FT DISULFID 81 167 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 167 183 PROBABLE.
 FT DISULFID 111 220 PROBABLE.
 FT CONFLICT 13 13 P -> L (IN REF. 4).
 FT CONFLICT 71 71 M -> I (IN REF. 3).
 SQ SEQUENCE 297 AA: 33077 MW: AC5420F8B626BDD1 CRC64;

Query Match 100.0%; Score 243; DB 1; Length 297;
 Best Local Similarity 100.0%; Pred. No. 9.5e-25;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KISHFLKMSLNPIRAHPTIYININCEPANSEKNSPTQYCS 44
 DB 142 KISHFLKMSLNPIRAHPTIYININCEPANSEKNSPTQYCS 185

RESULT 2
 CD20_MOUSE STANDARD; PRT: 291 AA.
 AC P19437.
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE B-cell surface protein CD20 homolog (B-cell differentiation antigen
 DE LY-44)
 GN MS4A1 OR CD20 OR LY-44 OR MS4A2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89067519; PubMed=2461992;
 RA Tedder T.F., Klejman G., Distche C.M., Adler D.A., Schlossman S.F.,
 RA Saito H.;
 RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte
 RT differentiation antigen, homologous to the human B1 (CD20) antigen,
 RT and localization of the gene to chromosome 19.";
 RL J. Immunol. 141:4388-4394(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;

RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner T., Gissi C., King B., Kochiwa H.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Schriml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 RA Blake J., Boffelli P., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
 CC B-CELL ACTIVATION AND PROLIFERATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PIR: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M62541; AAA37394.1; -
 DR EMBL: AK017903; BAB30996.1; -
 DR PIR: A30558; A30558.
 DR MGD: MGI:88321; MS4A2.
 KW B-cell; Transmembrane; Phosphorylation.
 FT DOMAIN 1 44 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 45 65 POTENTIAL.
 FT TRANSMEM 69 89 POTENTIAL.
 FT TRANSMEM 112 132 POTENTIAL.
 FT TRANSMEM 183 203 POTENTIAL.
 FT DOMAIN 204 291 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 291 AA: 31958 MW: DF478BCD2C5C16FC CRC64;

Query Match 65.8%; Score 160; DB 1; Length 291;
 Best Local Similarity 65.1%; Pred. No. 7.9e-14;
 Matches 28; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 2 ISHFLKMSLNPIRAHPTIYININCEPANSEKNSPTQYCS 44
 DB 137 ISHFLKMSLNPIRAHPTIYININCEPANSEKNSPTQYCS 179

RESULT 3
 YLM5_CAEEL STANDARD; PRT: 443 AA.
 ID YLM5_CAEEL
 AC P34379.
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein D2007.5 in chromosome III.
 GN D2007.5
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;

ID	SREP_PENCH	STANDARD:	PR1:	532 AA.
AC	O92259;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	GATA factor SREP.			
GN	SREP.			
OS	Penicillium chrysogenum.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.			
OX	NCBI_TaxID=5076;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Q176;			
RX	MEDLINE=97169143; PubMed=9016950;			
RA	Haas H., Angermayr K., Stoefli G.;			
RT	"Molecular analysis of a Penicillium chrysogenum GATA factor encoding			
RT	gene (srep) exhibiting significant homology to the Ustilago maydis			
RT	urs1 gene.";			
RL	Gene 184:33-37(1997).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- SIMILARITY: CONTAINS 2 GATA-TYPE ZINC FINGERS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U48414; AAC49628.1; -.			
DR	HSSP; P17679; 1GNF.			
DR	TRANSFAC; T02823; -.			
DR	InterPro; IPR000679; Znf_GATA.			
DR	Pfam; PF00320; GATA_2.			
DR	PRINTS; PR00619; GATAZFINGER.			
DR	SMART; SM00401; Znf_GATA; 2.			
DR	PROSITE; PS00344; GATA_ZN_FINGER_1; 2.			
DR	PROSITE; PS50114; GATA_ZN_FINGER_2; 2.			
KW	DNA-binding; zinc-finger; Transcription regulation; Nuclear protein.			
FT	DNA-binding 94 118 GATA-TYPE 1.			
FT	FT ZN-FING 144 262 GATA-TYPE 2.			
FT	DOMAIN 144 147 POLY-PRO.			
SO	SEQUENCE 532 AA; 56961 MW; D76EBE803C49EB0 CRC64;			
Query Match		21.4%;	Score 52;	DB 1; Length 532;
Best Local Similarity		46.2%;	Pred. No. 26;	
Matches 12;	Conservative 3;	Mismatches 11;	Indels 0;	Gaps 0;
OY	15 RAHRYIYINCEPANPSEKNSPTQ 40			
	I: I I I I I I I I I I I I I I I			
Db	59 RSGPGSNQMGHNETPSEKTPSSQ 84			
RESULT 9				
SC02_NEUCR	SC02_NEUCR	STANDARD:	PRT;	650 AA.
ID	SC02_NEUCR			
AC	Q01277;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Sulfur controller-2 (SCON2).			
GN	SCON-2.			
OS	Neurospora crassa.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			
OC	Sordariales; Sordariaceae; Neurospora.			
OX	NCBI_TaxID=5141;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=74-OR23-1A;			
RX	MEDLINE=95241499; PubMed=7724564;			
RA	Kumar A., Palella J.V.;			

```
RT "The sulfur controller-2 negative regulatory gene of Neurospora
RT crissa encodes a protein with beta-transducin repeats." ;
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995) .
CC -I- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -I- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -I- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS) .
CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -I- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
-----
CC
DR EMBL, U17251; AAA68968.1; -.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR Pfam: PF00646; F-box; 1.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 4.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
RW Transcription regulation / Repeat; WD repeat.
FT DOMAIN 124 170 F-Box.
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

Query Match 21.4%; Score 52; DB 1; Length 650;
Best Local Similarity 33.3%; Pred.No. 33;
Matches 11; Conservative 7; Mismatches 13; Indels 2; Gaps 1.;

QY 13 PIRATPYINYINCEPAN--PSKNSPSTQICY 43
   1:| | | | :| | :|::|||
DB 13 FLREHPSIVAPIGRGNGQTARAENPNISKYC 45

RESULT 10
B2MG_PITIR
ID B2MG_PITIR STANDARD; PRT; 119 AA.
AC 077531;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Beta-2-microglobulin precursor.
GN B2M.
OS Pithhecia irrorata (Bald-faced saki) .
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mamalia; Eutheria; Primates; Platyrrhini; Cebidae; Pitheciinae;
OC Pithecia.
OX NCBI_TaxID=30598;
RX MEDLINE=98298008; PubMed=9634477;
RA Canavez F.C., Ladaesky J.J., Muniz J.A.P.C., Seanez H.N., Parham P.;
RL Immunogenetics 48:133-140(1998) .
```


DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
 DE (ARS).
 OS Hemieutrotus pulcherrimus (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Strongylocentrotidae;
 OC Hemieutrotus.
 NCBI_TaxID=7650;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pluteus;
 RX MEDLINE=89030699; Pubmed=3181160;
 RA Sasaki H., Yamada K., Akasaka H., Suzuki K., Saito A., Sato M.,
 RA Shimada H.;
 RT "cDNA cloning, nucleotide sequence and expression of the gene for
 RT arylsulfatase in the sea urchin (Hemieutrotus pulcherrimus)
 RT embryyo.";
 RL Eur. J. Biochem. 177:9-13(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90092130; Pubmed=2598936;
 RA Yamada K., Akasaka K., Shimada H.;
 RT "Structure of sea-urchin arylsulfatase gene.";
 RL Eur. J. Biochem. 186:405-410(1989).
 CC -1- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE EXTRACELLULAR
 CC MATRICES INVOLVED IN CELL MOVEMENT DURING MORPHOGENESIS.
 CC -1- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
 CC -1- SUBCELLULAR LOCATION: IN BOTH THE CYTOPLASM AND THE EXTRACELLULAR
 CC MATRICES.
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17015; CAA34881.1; -.
 DR PIR: S01793; S01793.
 DR PIR: S07089; S07089.
 DR HSSP: P15289; IAKK.
 DR InterPro: IPR000917; Sulfatase.
 DR Pfam: PF00884; Sulfatase.1.
 DR PROSITE: PS00523; SULFATASE.1; 1.
 DR PROSITE: PS00149; SULFATASE.2; 1.
 KM Hydroxylase; signal; Glycoprotein; Extracellular matrix.
 FT STGNAL 1 20
 FT CHAIN 21 551
 FT MOD_RES 21 21
 FT MOD_RES 100 100
 FT ACT_SITE 158 158
 FT CARBOHYD 164 164
 FT CARBOHYD 213 213
 FT CARBOHYD 296 296
 SO SEQUENCE 551 AA; 60952 MW; 54C1AAC14D6710C9 CRC64;
 Query Match 21.2%; Score 51.5; DB 1; Length 551;
 Best Local Similarity 24.4%; Pred. No. 32;
 Matches 10; Conservative 10; Mismatches 20; Indels 1; Gaps 1;
 Oy 3 SHFLKMSLNIRAHPTIYNNEPANPSKNSPSTOYCY 43
 Db 170 AHLPEHNGHFDVGHNLPTNSWSCDDTG-LHKDPDSORCY 209
 RESULT 14
 VGF_BPG4 STANDARD; PRT; 426 AA.
 AC P03642;
 DT 21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Capsid protein (F protein) (CPF).
 GN F.
 OS Bacteriophage G4.
 OC Viruses; ssDNA viruses; Microviridae; Microvirus.
 NCBI_TaxID=10843;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=79053264; Pubmed=714153;
 RA Godson G.N., Barrell B.G., Staden R., Fiddes J.C.;
 RT "Nucleotide sequence of bacteriophage G4 DNA.";
 RL Nature 276:236-247(1978).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS), AND REVISIONS TO 35-36;
 RP 68 AND 281.
 RX MEDLINE=96217890; Pubmed=8642594;
 RA McKenna R., Bowman B.R., Liag L.L., Rossmann M.G., Fane B.A.;
 RT "Atomic structure of the degraded procapsid particle of the
 RT bacteriophage G4: induced structural changes in the presence of
 RT calcium ions and functional implications.";
 RL J. Mol. Biol. 256:736-750(1996).
 CC -1- FUNCTION: F PROTEIN IS THE MAJOR CAPSID COMPONENT.
 CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
 CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V00657; CAA24019.1; -.
 DR PIR: A04250; ZEPBPG4.
 DR PDB: 1GF; 03-APR-96.
 DR InterPro: IPR003514; Capsid_F.
 DR Pfam: PF02305; Phage_F.1.
 KM Coat protein; 3D-structure.
 FT INT_MET 0 0
 FT SEQUENCE 426 AA; 48572 MW; 539013960DAF51E CRC64;
 Query Match 21.0%; Score 51; DB 1; Length 426;
 Best Local Similarity 35.3%; Pred. No. 27;
 Matches 16; Conservative 3; Mismatches 8; Indels 22; Gaps 4;
 Oy 1 KISHFLKMSLNIRAHPTIYN--CEP-----ANPSEKNSPSTOY 41
 Db 118 KVPKFL-----HGGYININYNFKPPMSDDLTYANPS--NMPSEDY 156
 RESULT 15
 K6PF_METUA
 ID K6PF_METUA STANDARD; PRT; 462 AA.
 AC Q58999;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE ADP-specific phosphofructokinase (EC 2.7.1.146) (ADP-dependent
 DE phosphofructokinase) (ADP-Pfk).
 GN PFKC OR MJ1604.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; Pubmed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstein K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Ullrich J., Kelley J.M., Peterson J.D., Sadow P.M., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*.";
 RT Science 273:1058-1073(1996).
 RN [2]
 RP CHARACTERIZATION.
 RX PubMed-117127273;
 RA Verhees C.H., Tulinaga J.E., Kengen S.W.M., Stams A.J.M.,
 RA van der Oost J., de Vos W.M.;
 RT "ADP-dependent phosphofructokinases in mesophilic and thermophilic
 RT methanogenic archaea.";
 RL J. Bacteriol. 183:7145-7153(2001).
 CC -1- FUNCTION: Catalyzes the phosphorylation of fructose 6-phosphate to
 CC fructose 1,6-bisphosphate using ADP as the phosphate donor.
 CC -1- CATALYTIC ACTIVITY: ADP + D-fructose 6-phosphate = AMP + D-
 CC fructose 1,6-bisphosphate.
 CC -1- PATHWAY: Glycolysis, modified hyperthermophilic version.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PFKC FAMILY OF CARBOHYDRATE KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL; U67600; AAB99627.1; -;
 DR TIGR; M1604; -;
 KW Transferase; Kinase; Glycolysis; Complete proteome.
 SQ SEQUENCE 462 AA; 5361 MW; D9108B60DBA104 CRC64;

Query Match 21.0%; Score 51; DB 1; Length 462;
 Best Local Similarity 32.1%; Pred. No. 30;
 Matches 9; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Oy 7 KMESLNFRAHTPYINYNCEPANPEK 34
 |::|::|::|::|::|::|::|::|::|
 Db 338 KFKNLEVVQVHTIYILFVCRADNPLSK 365

Search completed: March 19, 2003, 17:21:38
 Job time : 10.5 secs

AC Q978H7;

```
* Query Match      27.6%   Score 67 : DB 17; Length 316;  
Best Local Similarity 35.9% ; Pred. NO 0.3;  
Matches    14; Conservative    7; Mismatches    18; Indels    0; Gaps    0;
```

	Query Match	25.3%	Score 61.5;	DB 4;	Length 580;
	Best Local Similarity	35.5%	Pred. No. 3.3;		
	Matches	11;	Conservative	5;	Mismatches 14; Indels 1; Gaps 1;
OY	14 IRAHPYIINTCEPANPESEKNSPSTOYCYS	44			
	:::: :::	:::: :::			
b	203 LKAHQPY-GVQACVPCCGPGTNNRKHSLCYN	232			

RP SEQUENCE FROM N.A.
RA Chaur R., Shinn P., Brooks S., Buehler E., Chao O., Johnson-Hopson C.,
RA Khan S., Kim C., Altfelt H., Bel B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharasy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southick A.,
RA Thavert A., Tolmuni M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologidis A., Ecker J.
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC032279; AAF79257.1; -
DS SEQUENCE 702 AA; 80578 MW; 425B0F22B75EC786 CRC64;

Query March	25.1%	Score 61;	DB 10;	Length 702;
Best Local Similarity	35.5%	Pred. NO. 4.7;		
Matches 11;	Conservative	8;	Mismatches	12;
			Indels	0;
			Gaps	0;

OY	14	IRAHPTYININCEPANPEKNSPSTQCYCS	44
		: ::: ::	
Db	624	VRAHPTMSNLFCSLMFNEVERFTRDOLSLFA	654

RESULT 5	
027732	
ID	027732
AC	027732;
DT	01-NOV-1996 (TREMBlrel. 01, Created)
DT	01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE	Copper amine oxidase (EC 1.4.3.6).
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5833;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FCO27;
RA	Flores M.V.C.;
RT	"Molecular Biology of Pyrimidine Biosynthesis in the Human Malarial Parasite Plasmodium falciparum.";

RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCO27;
 RA Flores M.V.C., O'Sullivan W.J., Stewart T.S.;
 RT "Characterisation of the Cardamoyl Phosphate Synthetase Gene from
 RL Plasmodium falciparum.";
 RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
 CC H(2)O(2).
 CC -1- COFACTOR: COPPER AND TOPAQUINONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE COPPER/TOPAQUINONE OXIDASE FAMILY.
 DR EMBL: L32150; AAA29522.1; -.
 DR HSSP: P00968; 1A9X.
 DR InterPro: IPR000901; CPase.
 DR InterPro: IPR002474; CPase_sm_chain.
 DR InterPro: IPR00317; CPS_GATase.
 DR InterPro: IPR000269; CuNH_oxidase.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000991; GATase_1.
 DR InterPro: IPR004362; MGS_like.
 DR Pfam: PF00289; CPase_L_chain; 2.
 DR Pfam: PF02786; CPase_L_D2; 2.
 DR Pfam: PF02787; CPase_L_D3; 1.
 DR Pfam: PF00988; CPase_sm_chain; 1.
 DR Pfam: PF00117; GATase; 1.
 DR PRINTS: PRO0098; CPASE.
 DR PRINTS: PRO0099; CPSEGATASE.
 DR PRINTS: PRO0096; GATASE.
 DR PROSITE: PS01164; COPPER_AMINE_OXID_1; 1.
 DR PROSITE: PS00866; CPASE_1; 1.
 DR PROSITE: PS00867; CPASE_2; UNKNOWN_2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR COPPER; Ligase; Oxidoreductase; TPQ.
 KW SEQUENCE 2391 AA; 275674 MW; 426F3C3E612FEE2 CRC64;
 SQ
 Query Match 24.9%; Score 60.5; DB 5; Length 2391;
 Best Local Similarity 34.3%; Pred. No. 20;
 Matches 12; Conservative 10; Mismatches 12; Indels 1; Gaps 1;
 QY 2 ISHFLKMSLNFIKRAHPYI-NIYNCEPANSEKN 35
 Db 317 INHTLRDKMNLITSESEYTKDLNCFNSGSDKN 351
 ID Q9CLP3 PRELIMINARY; PRT; 303 AA.
 AC Q9CLP3;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical protein PM1174.
 GN PM1174.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella
 OC NCBL_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AE006157; AAK03258.1; -.
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 303 AA; 35676 MW; BAF7BC8C7563E921 CRC64;
 SQ
 Query Match 24.1%; Score 58.5; DB 16; Length 303;
 Best Local Similarity 31.2%; Pred. No. 4.5;
 Matches 10; Conservative 11; Mismatches 10; Indels 1; Gaps 1;

QY 6 LKMSLNFIKRAHPYI-NIYNCEPANSEKN 37
 Db 202 IDLEALSTLQOHY-HVNFYISIPMSPLSKHP 232
 ID 055208 PRELIMINARY; PRT; 194 AA.
 AC 055208;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Transcription factor Figa.
 GN FIGLA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBL_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98090049; PubMed=9362457;
 RA Liang L., Soyai S.M., Dean J.;
 RT "Figa1pha, a germ cell specific transcription factor involved in the
 RT coordinate expression of the zona pellucida genes.";
 RL Development 124:4939-4947(1997).
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 DR EMBL: U91840; AAB97139.1; -.
 DR HSSP: P25912; 1HLO.
 DR MGD: MGI:1349421; Figa.
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
 KW DNA-binding.
 SQ SEQUENCE 194 AA; 21529 MW; B12137E37AF22723 CRC64;
 QY 7 KMESLNFIKRAHPYI-NIYNCEPANSEKN 40
 Db 94 KPSKVDILKATETITGLCYLEAKVSEKOSPEEQ 129
 ID 08VWJ8 PRELIMINARY; PRT; 1210 AA.
 AC 08VWJ8;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Putative polyprotein.
 GN OSUNBA0023119.20 OR OSUNBA0079B05.3.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartioideae; Oryzaceae; Oryza.
 OC NCBL_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sasaki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC079037; AAL34947.1; -.
 DR EMBL: AC079179; AAL31648.1; -.
 DR InterPro: IPR005162; Retrotrans_gag.
 DR InterPro: IPR001584; Rve.

DR InterPro: IPR000477; RVTSE.
 DR Pfam: PF03732; Retrotrans_gag; 1.
 DR Pfam: PF00665; rve; 1.
 DR Pfam: PF00078; rvc; 1.
 KW Polypeptide.
 SQ SEQUENCE 1210 AA; 133656 MW; 51687F890214E71F CRC64;

Query Match 23.5%; Score 57; DB 10; Length 1210;
 Best Local Similarity 23.1%; Pred. No. 30;
 Matches 15; Conservative 8; Mismatches 14; Indels 28; Gaps 2;

OY 6 LKESINFRAR-----TRYINYNCEPANPSE---KNSP 37
 DB 1038 LKTKFTNLIKKGDSNIEELPAVLNMRTPSRATGETPFPLVYGAVALPSLTLKSPR 1097
 OY 38 STQYC 42
 DB 1098 ATMYC 1102

RESULT 9
 O28859 PRELIMINARY; PRT: 2425 AA.
 AC O28859:
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Surface layer protein B (SLCB-2).
 GN AF1413.
 OS Archaeoglobus fulgidus.
 UC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RA MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.D., Kierliavage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirtness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Arllich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Weese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE001006; AAB89834.1; -.
 DR TIGR: AF1413; -.
 DR InterPro: IPR003961; FN_III.
 DR Pfam: PF00041; fn3; 1.
 DR PRINTS: PRO1656; VACCYTOTOXIN.
 DR SMART: SM00060; FN3; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 2425 AA; 266571 MW; 97E3856843226E37 CRC64;

Query Match 23.3%; Score 56.5; DB 17; Length 2425;
 Best Local Similarity 30.0%; Pred. No. 73;
 Matches 15; Conservative 9; Mismatches 19; Indels 7; Gaps 2;

OY 1 KISHFLKMSLNFIRAHPIYINYN---CEPANPSE---KNSPSQYCY 43
 DB 1764 KISSFLANATEFVNSTSPEVEYVNDLIEVLNQAETALRNPGMDPIFY 1813

RESULT 10
 O9SST1 PRELIMINARY; PRT: 711 AA.
 AC O9SST1:

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE Hypothetical 81.3 kDa protein.
 GN AT4G09630.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McCombie R.W., Spiegel L.A., Huang E.N., Nascimento L.U.,
 RA de la Bastide M., Vil D.M., Preston R.R., Matero A., Rodriguez M.,
 RA O'Shaughnessy A., Rodriguez M., Shekher M., Schuit K., See L.H.,
 RA Swaby I., Habermann K., Dedhia N.N., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 RA Shekher M., Schuit K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL161831; CAB82130.1; -.
 DR EMBL: AL161515; CAB78086.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 711 AA; 81270 MW; E49BBE6FD92AF5A CRC64;

Query Match 23.0%; Score 56; DB 10; Length 711;
 Best Local Similarity 32.3%; Pred. No. 24;
 Matches 10; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

OY 14 IRAHPIYINYNCEPANPSEKNSPSQYCY 44
 DB 632 VREHTPMNSLFCMLPNEVEKTPRDLSFA 662

RESULT 11
 O73781 PRELIMINARY; PRT: 299 AA.
 AC O73781:
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Fork head domain protein FKD5.
 GN FOXB1.1 OR FKD5.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98350246; PubMed=9683740;
 RA Odenthead J., Nusslein-Volhard C.;
 RT "fork head domain genes in zebrafish.";
 RL Dev. Genes Evol. 208:245-258(1998).
 DR EMBL: AF052248; AAC06365.1; -.
 DR HSSP: Q63245; 2HRH.
 DR ZFIN: ZDB-GENE-990415-77; foxb1.1.
 DR InterPro: IPR001766; TF_Fork_head.
 DR Pfam: PF00250; Fork_head; 1.
 DR PRINTS: PR00053; FORKHEAD.
 DR ProDom: PD000425; TF_Fork_head; 1.
 DR SMART: SM00339; FH; 1.
 DR PROSITE: PS00657; FORK_HEAD_1; 1.
 DR PROSITE: PS00658; FORK_HEAD_2; UNKNOWN_1.
 DR PROSITE: PS00039; FORK_HEAD_3; 1.
 SQ SEQUENCE 299 AA; 33219 MW; A2C731F218EA9EAA CRC64;

Query Match 22.8%; Score 55.5; DB 13; Length 299;

Best Local Similarity 27.5%; Pred. No. 12;
Matches 11; Conservative 9; Mismatches 17; Indels 3; Gaps 1;

QY 3 SHFLKMS---LNFIRAHPTIYINCEPANPSEKNSPST 39
DB 125 AHYLOOAKLRMTALGTHLPQMTSYNLSVTPQSTLNPRTL 164

RESULT 12

ID 025384 PRELIMINARY; PRT; 551 AA.

AC 025384;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Arylsulfatase.

GN ARS.

OS Hemicentrotus pulcherrimus (Sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;

OC Hemicentrotus.

OC NCBI_Taxid=7650;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-SPERM;

RA Yamada K.;

RL Submitted (SEP-1989) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 1-105 FROM N.A.

RC TISSUE-SPERM;

RA MEDLINE-90092130; PubMed-2598936;

RA Yamada K., Akasaka K., Shimada H.;

RT "Structure of sea-urchin arylsulfatase gene.,"

RL Eur. J. Biochem. 186:405-410(1989).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE-SPERM;

RA Shimada H.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE-SPERM;

RA Akasaka K., Sakamoto N., Yamamoto T., Morokuma J., Fujikawa N.,

RA Takata K., Eguchi S., Shimada H.;

RT "Corrected structure of the 5'flanking region of arylsulfatase gene of

RT the sea urchin, Hemicentrotus pulcherrimus.,"

RL Dev. Growth Differ. 36:633-636(1994).

DR EMBL: X16679; CAA34667.1; -.

DR EMBL: X16680; CAA34667.1; JOINED.

DR EMBL: X16681; CAA34667.1; JOINED.

DR EMBL: X16682; CAA34667.1; JOINED.

DR EMBL: X16683; CAA34667.1; JOINED.

DR EMBL: X16684; CAA34667.1; JOINED.

DR HSSP: P15289; IAUK.

DR InterPro: IPR000917; Sulfatase.

DR Pfam: PF00884; Sulfatase.1.

DR PROSITE: PS00523; SULFATASE.1; 1.

DR PROSITE: PS00149; SULFATASE.2; 1.

SQ SEQUENCE 551 AA; 60943 MW; 544315A480E12C1 CRC64;

Query Match 22.8%; Score 55.5; DB 5; Length 551;

Best Local Similarity 26.8%; Pred. No. 22;

Matches 11; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

QY 3 SHFLKMSLNFIRAHPTIYINCEPANPSEKNSPSTQYCY 43

DB 170 AHPLENGHGFVGHNLPTNSWSCDDTG-LHKDFPDTQCY 209

RESULT 13

ID 044131 PRELIMINARY; PRT; 2150 AA.

AC 044131;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE C08G9.2 protein.

GN C08G9.2.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OC NCBI_Taxid=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA MEDLINE-94150718; PubMed-7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,

RA Snelson N., Smith A., Sonhammer E., Staden R., Sulston J.,

RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RT elegans.,"

RL Nature 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Geisel C., Stellyes L.;

RT "The sequence of C. elegans cosmid C08G9.,"

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Waterston R.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF036687; AAB88311.1; -.

DR HSSP: P10646; IADZ.

DR InterPro: IPR004094; Antistatin.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR002223; Kunitz_BPTI.

DR InterPro: IPR000716; Thyroglobulin_1.

DR InterPro: IPR002899; WAP.

DR InterPro: IPR000822; ZnF_C2H2.

DR Pfam: PF02822; Antistatin; 6.

DR Pfam: PF00014; Kunitz_BPTI; 1.

DR Pfam: PF00086; thyroglobulin_1; 6.

DR Pfam: PF00095; wap; 9.

DR PRINTS: PR00003; ADISULPHOCORE.

DR PRINTS: PR00759; BASICPTASE.

DR PRODOM: PD000222; Kunitz_BPTI; 1.

DR SMART: SM001224; WAP; 4.

DR SMART: SM00131; KU; 1.

DR SMART: SM00211; TV; 6.

DR SMART: SM00217; WAP; 9.

DR SMART: SM00289; WRI; 6.

DR PROSITE: PS00317; 4_DISULFIDE_CORE; 9.

DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.

DR PROSITE: PS50279; BPTI_KUNITZ_2; 1.

DR PROSITE: PS00022; EGF_1; UNKNOWN_1.

DR PROSITE: PS00484; THYROGLOBULIN_1; UNKNOWN_1.

DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

KW Serine protease inhibitor

SQ SEQUENCE 2150 AA; 234131 MW; 4ED66BD669C7523C CRC64;

Query Match 22.8%; Score 55.5; DB 5; Length 2150;

Best Local Similarity 41.2%; Pred. No. 89;

Matches 14; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 11 LNFIRAHPTIYINCEPANPSEKNSPSTQYCY 43

DB 111 LNFIRAHPTIYINCEPANPSEKNSPSTQYCY 43

```

DB 610 LNFCEPGEPIYISSIGNAESCETKDECPSTHCH 643
RESULT 14
O9JUT4
AC O9JUT4: PRELIMINARY: PRT: 149 AA.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE RNA polymerase (Fragment).
OS Gnetum guenon.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales;
OX NCBI_TaxID=3382;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99406856; PubMed=10473771;
RA Samigullin T., Martin W., Trolitsky A., Antonov A.;
RT "Molecular data from the chloroplast rpoC1 gene suggest a deep and
RT distinct dichotomy of contemporary spermatophytes into two
RT monophylums: gymnosperms (including Gnetales) and angiosperms.";
RU J. Mol. Evol. 49:310-315(1999).
DR EMBL, AJ012567; CAB56787.1; -;
DR InterPro: IPR000722; RNA_pol_A.
DR Pfam: PF00623; RNA_pol_A; 1.
KW Chloroplast.
FT NON_TER 1 1
FT SEQUENCE 149 AA: 17137 MW: 09103F3C78D4DE52 CRC64;
SQ
Query Match 22.6%; Score 55; DB 8; Length 149;
Best Local Similarity 34.8%; Pred. No. 6.7;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
Qy 21 INIYCEPANPSEKNSPSTOYCY 43
DB 55 ISOFCRCPSPNPTKFLPEVNLCF 77
RESULT 15
O9LOV8
AC O9LOV8: PRELIMINARY: PRT: 380 AA.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F10B6.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cho Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,
RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
RA Hansen N.F., Huitzer L., Kremenetskaia I., Lenz C., Li J., Liu S.,
RA Luros S., Rowley D., Schwartz J., Tortum M., Vysotskaia V., Yu G.,
RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
RT 1.";
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [4]

```

```

RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howling B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thayerl A., Tortum M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AC006917; AAF9222.1; -;
DR InterPro: IPR004877; CYL_B561.
DR Pfam: PF03188; Cytochrome_B561; 1.
SQ SEQUENCE 380 AA: 42597 MW: E32C44ABF67109F8 CRC64;
Query Match 22.6%; Score 55; DB 10; Length 380;
Best Local Similarity 27.8%; Pred. No. 18;
Matches 10; Conservative 9; Mismatches 17; Indels 0; Gaps 0;
Qy 6 LKMSLNFTRAHPTPIYINCEPANPSEKNSPSTOY 41
DB 197 LKLVNVSFEYIHFLYVSMYTCQIKKKNYTPMPTY 232
Search completed: March 19, 2003, 17:21:11
Job time : 50 secs

```